Synopsis zur kumulativen Dissertation
zur Erlangung des Doktorgrades (Dr. rer. nat.)
am Fachbereich 06 der Justus-Liebig-Universität Gießen
Mai 2023

Mechanisms of Approximate Numerosity Processing: Approximate Number Tasks in the Haptic and Visual Modality

Vorgelegt von:
Marco Carlo Ziegler

Prüfungsgremium

Gutachter und Erstbetreuer: Prof. Dr. Martin Kersting, FB 06 Justus-Liebig-Universität Gießen

Zweitgutachter: Prof. Dr. Jutta Billino (apl.), FB 06 Justus-Liebig-Universität Gießen

Zweitbetreuer: Prof. Dr. Knut Drewing (apl.), FB 06 Justus-Liebig-Universität Gießen

Externer Prüfungsbeisitzender: Prof. Dr. Matthias Vogel, FB 04 Justus-Liebig-Universität Gießen

Drittgutachter: Prof. Dr. Avishai Henik (Ben-Gurion University of Negev)

Danksagung

Eine Arbeit zu einem Thema schreiben zu dürfen, das man liebt, ist ein Privileg. Wenn diese Arbeit zudem noch von vielen Menschen unterstützt wird, ist dies ein zusätzliches Glück.

Ich danke meinen Doktorvätern Prof. Dr. Martin Kersting und Prof. Dr. Knut Drewing dafür, dass sie mich während meiner gesamten Promotionszeit unterstützt und bestärkt haben, zu einem meiner liebsten Themen zu promovieren. Martin gebührt mein besonderer Dank dafür, dass er mir in seiner Abteilung, der Psychologischen Diagnostik, die Promotion ermöglicht hat. Ich danke ihm für fachliche und überfachliche Beratung, für Vertrauen und dafür, dass er immer eine offene Tür und ein offenes Ohr für mich hatte. Knut danke ich, neben der fachlichen und überfachlichen Beratung insbesondere dafür, dass er mit mir intensiv fachliche Themen reflektiert hat, mich rund um den Publikationsprozess der Studien gestützt und unterstützt und mir dabei viel beigebracht hat. Vielen Dank an euch beide, ich hätte mir keine besseren Doktorväter wünschen können. Ich danke weiterhin Prof. Dr. Jutta Billino dafür, dass sie diese Arbeit begutachtet hat. Prof. Dr. Matthias Vogel danke ich dafür, dass er sich bereit erklärt hat, als externer Beisitzender die Prüfungskommission zu vervollständigen.

Ein besonders liebevoller Dank geht an Manuela Schirle für das viele Korrekturlesen meiner Artikel und für die Vorschläge, die sie dabei eingebracht hat. Ich kann dir kaum genug dafür danken- tausend Dank an dich! Ich danke meinen vielen tollen Kolleg*innen aus dem Fachbereich 06, insbesondere allen Kolleg*innen der Psychologischen Diagnostik und des HapLab, für die jahrelange tolle gemeinsame Arbeit an vielen unterschiedlichen Themen. Außerdem geht ein Dank an meine (ehem.) Doktoranden- und Bandkollegen Dr. Aaron Zöller, Frieder Hartmann und Marcel Schepko, die während der Promotionszeit auch für ein aufregendes außeruniversitäres Leben gesorgt haben. Ein weiterer Dank geht an alle studentischen Hilfskräfte, die geholfen haben, Daten zu erheben, um die Projekte zu realisieren. Ein großes Dankeschön an Luisa Stricker, die kontinuierlich an meinen „ANS"-Projekten beteiligt und engagiert war.

Das nächste große Dankeschön geht an meine gesamte Familie und an meine langjährigen Freunde. Freie Zeit war während der Promotion immer eine knappe Ressource. Ich bin froh darüber, dass feste Bande wartungsarm sind und die immer zu kurze gemeinsame Zeit intensiv und schön war und ist.

Ich danke meiner Mutter Sandra Ziegler und meinen Großeltern, Renate und Manfred Jag, für ihre unendliche Liebe, Unterstützung und Fürsorglichkeit über alle meine bisherigen Lebensjahre.

Viele weitere Menschen sollten hier erwähnt werden, weil sie mich unterstützt, gefördert oder inspiriert haben. Ich muss mich an dieser Stelle aber auf ein einfaches Dankeschön beschränken, Dankeschön!

Addendum (22.12.2023): I would like to thank Prof. Dr. Avishai Henik for reviewing my dissertation and for joining the examination committee.

Abstract

The perception and processing of number is an essential ability of humans and animals to interact with their environment. The Approximate Number System (ANS) is assumed to be a part of the perception and cognitive processing of distinct numerical magnitude without counting. Non-symbolic perception of numerosity in an imprecise manner allows advantageous real-life decision making. A modern example from everyday life might be that if you want to get to a checkout faster, it makes sense to join a queue with fewer people. This decision, based on the number of people in the line, can be made in a fraction of a second without actually counting the number of people. The Approximate Number System is thought to allow extracting the approximate number of items in scenes, such as the number of people in a queue. One popular assumption about the ANS is that number is an abstract representation and not tied to a sensory modality, since stimulus-related features that go along with discrete entity (like the occupied area) are removed. Although widely accepted, this assumption of an amodal percept lacks on profound empirical evidence. Only very few studies have explored approximate numerosity in other modalities than the visual. Furthermore, different factors regarding the methodology for assessing ANS-acuity and recent findings indicate that other influences (e.g., spatial attributes of visual stimuli) also affect a numerosity percept, questioning the theoretical ANS assumptions. This work contributes in forthcoming and understanding the nature of the ANS by utilizing intra- and cross-modal paradigms adapted to the visual and haptic modality.

In Study 1, we used a dot comparison task in both the visual modality and the haptic modality to measure the non-symbolic numerical discrimination ability of 67 participants. We found that a) besides number, participants were also influenced by spatial factors in their performance, and b) even though the performance was ratio-dependent, indicating ANS involvement, the performance between the modalities was not correlated.

In Study 2, we investigated in 50 participants whether their discrimination performance was affected by spatial information of numerical stimuli in an approximate number-matching task. We let participants perceive numerosity from haptic source stimuli and then let them match it to one of two visually presented dot patterns. We found that the spatial configurations of the (dot) stimuli in the haptic source modality significantly affected the numerosity matching performance in the visual modality.

The findings of both studies are in conflict with the strong assumptions of the so-called “direct” ANS model, which is currently state-of-the-art. Due to our results, there is reason to widen the perspective of how the percept of numerosity arises and to consider competing ANS perspectives or alternative approaches for empirical testing. Furthermore, we emphasize that the measurement of ANS-acuity requires better standardization and sensitivity for specific implementations in the future.

Inhalt

I Synopsis

1 Introduction

What abilities are so essential and useful to human cognition that nature would not leave it to chance whether or not you acquire them? It must be a set of abilities, which enables you to interact with the environment in such a way as to ensure survival and thriving throughout your life span. Spelke and Kinzler (2007) propose that human cognition roots in flexible but distinct core systems, each having a signature contribution to the functioning of human beings. Different domains have been considered candidates for core systems or core knowledge, among them a core understanding for physics, actions, food, geometry, language, and number (Barner & Baron, 2016; Feigenson et al., 2004; Shutts et al., 2009; Spelke & Kinzler, 2007; Strickland, 2017; Vallortigara, 2012). According to Spelke (2022), "number" is one of the domains that has been widely studied, and the idea of an innate number sense has probably become the most prominent theory in the field of numerical cognition (Dehaene, 2011; Lourenco & Aulet, 2022). The "number sense theory" is the foundation of what has developed into a "mainstream" theory (Leibovich, Katzin, Harel, et al., 2017) on how humans perceive and process number: The two core systems theory (Dehaene, 2011; Feigenson et al., 2004; Hyde, 2011; Leibovich, Katzin, Harel, et al., 2017).

1.1 Two Core Systems of Number

Feigenson et al. (2004) distinguish two distinct systems of numerical representation in humans. One system allows for a fast and precise processing of smaller numerosities (< 4) and another one processes larger numerosities in an imprecise and fuzzy manner (Dehaene, 2011; Feigenson et al., 2004; Hyde, 2011; Mou & vanMarle, 2014).

1.1.1 Precise Representations of Distinct Individuals

The first system is referred to as system for "precise representations of distinct individuals" (Feigenson et al., 2004) or sometimes as "object tracking system" or "parallel individuation system" (Hyde, 2011; Mou & vanMarle, 2014). Despite using different terms, all authors commonly use the behavioral observation as signature that humans can extract numerosity, defined as the "[...] the number of countable elements in a given group" (Nieder, 2016, p.366), from a small set of items accurately, almost instantaneously, and without counting (Feigenson et al., 2004; Hyde, 2011; Mou & vanMarle, 2014). This rapid enumeration process of small sets of elements without counting is also known as subitizing (Kaufman et al., 1949).

1.1.2 Approximate Number System and its Hallmarks

In distinction, the second numerical representation system processes larger numbers imprecisely without counting (Feigenson et al., 2004; Hyde, 2011). This system is referred to as the "approximate number system" (ANS; Brannon & Merritt, 2011; Dehaene, 2011; Feigenson et al., 2004; Hyde, 2011). The ANS is considered as a phylogenetically old cognitive system which different species share (Brannon & Merritt, 2011; Dehaene, 2011). Evidence for the processing of (approximate) numerical magnitudes has been found in primates such as chimpanzees [Pan troglodytes], gorillas [Gorilla gorilla], bonobos [Pan paniscus], and orangutans [Pongo pygmaeus] (Hanus & Call, 2007; Rumbaugh et al., 1987). In addition, various vertebrates such as chicks [Gallus gallus], anglefish [Pterophyllum scalare] and mosquitofish [Gambusia holbrooki] and also, invertebrates such as dune snails [Theba pisana] appear to process numerical magnitudes (Agrillo et al., 2008; Bisazza & Gatto, 2021; Gómez-Laplaza & Gerlai, 2011; Hanus & Call, 2007; Rugani et al., 2020) indicating the involvement of an ANS.

According to literature, the "approximate number system" can be characterized by different hallmarks or signatures: scalar variability, ratio dependence, approximate calculation and abstract number representation (Brannon & Merritt, 2011; Odic & Starr, 2018; Spelke & Kinzler, 2007).

Scalar variability can be best pictured in the context of an original experiment of Whalen et al. (1999), in which they instructed participants to press a key as fast as they could to match a certain presented number ranging between 7 to 25. They found that participants produced different variances in the number of key presses for the different numbers but a constant coefficient of variation, which reflects the relation between the target number and the amount of key presses. The participants' variance for 10 key presses is therefore half of the variance for 20 key presses.

Closely related is the *ratio dependency* as hallmark of the ANS. When humans compare numerical magnitudes, the difficulty and therefore the accuracy and speed of the comparison is determined by the ratio of the numerical magnitudes rather than their absolute cardinality. A comparison of 7:8 dots (ratio of 1.14) on a display is therefore more difficult than a comparison of 6:7 dots (ratio of 1.16) despite their absolute difference being equally one. Therefore, the ratio dependency follows Weber-Fechner's law[1] which "[...] states that the discriminability of two stimuli is linearly related to their ratio, equivalent to their distance on the logarithmic scale" (DeWind et al., 2015, p. 250). To illustrate this further, the numerosity of "8" and "16" are, according to this law, equally well distinguishable as "16" and "32". The ratio dependence signature is most commonly explained by the logarithmic number line model, which suggests that number is represented on an internal continuum, a mental number line, in which the numbers on the continuum are increasingly compressed for larger numerical magnitudes (Brannon & Merritt, 2011; Dehaene, 2011; Feigenson et al., 2004; Sasanguie et al., 2017). Furthermore, the neural activation of number is associated with an error (represented by a Gaussian curve) and therefore higher numbers increasingly overlap in their error margin, whereas more distant numerical magnitudes can be distinguished with better acuity (Brannon & Merritt, 2011; Dehaene, 2011; Feigenson et al., 2004; Sasanguie et al., 2017). Scalar variability and ratio dependency are closely related, since "Scalar variability is an instance of Weber's law" (Piazza, 2010, p. 542).

Another signature is that approximate numerical magnitudes can be actively manipulated and used for *arithmetic operations* (Brannon & Merritt, 2011; Spelke & Kinzler, 2007). For example, McCrink and colleagues demonstrated adding and subtracting approximate numerosities with dot patterns in children (McCrink & Wynn, 2004) and adults (McCrink et al., 2007). Furthermore, Qu et al. (2021) showed that children seem to be capable of approximate multiplication even before formally learning the concept of multiplication.

The last ANS signature is *amodality* of the numerosity representation (Spelke & Kinzler, 2007). The ANS amodality assumption is based on the idea that numerosity can be extracted from any kind of stimulus material as an abstract feature, regardless of its accompanying features or its source modality (Brannon & Merritt, 2011; Spelke & Kinzler, 2007). With the term "abstract" in the context of numerosity estimation, I agree with the explication: "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 [intraparietal sulcus], an area suggested to be dedicated to abstract number processing" (Gebuis et al., 2016, p.27). Gebuis et al. (2016) mention that this leans into the definition: "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, as cited in Gebuis et al., 2016). Complementarily, I concur with the following explication:

---

[1] In the following called Weber's law

> “[…] abstract representation in the present article is that neuronal populations that code numerical quantity are insensitive to the form of input in which the numerical information was presented (e.g., digits, verbal numbers, auditory, numerosity, etc.). In contrast, we define non-abstract representation as neuronal populations that code numerical quantity but are sensitive to the input in which the numbers were presented.” (Kadosh & Walsh, 2009, p. 314)

I think all three explications are helpful to describe different aspects of the term “abstract” in our context of non-symbolic approximate numerosity processing (but see Cantlon et al., 2009). As a consequence of these cited definitions, I suggest that an approximate “numerosity percept” (fuzzy by definition) extracted from different source modalities should be highly similar and robust across modalities and task implementations, and, following the “strong” interpretation of the ANS theory as stated above, should be based exclusively on number. In traditional ANS theories (or “number sense” theories), the creation of an abstract numerosity representation is described by a process with different stages (Dehaene, 2011; Dehaene & Changeux, 1993; Gebuis et al., 2016). In an initial “sensory” stage, the non-symbolic source stimulus (e.g., an array of dots) is received by a sensory modality. The non-symbolic numerical stimuli can vary in several characteristics that covary with number, such as the individual dot size of the set or the occupied area of the pattern (DeWind et al., 2015; Gebuis & Reynvoet, 2011). In the second stage, which is called the normalizing phase, these covariant stimulus characteristics are removed and a normalized signal for each dot remains. In the consecutive accumulation phase, the standardized “number signals” are summed and shape the final numerosity percept, which can also contain a degree of error (Gebuis et al., 2016; cf. Tokita et al., 2013). This process is typical for a “direct ANS model” (c.f. Qu et al., 2022) and is usually the default theory to explain how approximate numerosity is extracted from the environment and why a numerosity percept is assumed to be abstract (Dehaene, 2011; Halberda et al., 2008; Hyde, 2011; Spelke & Kinzler, 2007; Brannon & Merritt, 2011). However, in contrast to the other ANS hallmarks, such as the well-elaborated ratio dependency, the abstractness of the numerosity representation and its proposed mechanism have not been extensively studied. In fact, very little is known about approximate number processing in different modalities (Ziegler & Drewing, 2022).

1.1.3 The Approximate Number System and Mathematical Ability

The precision to discriminate or to represent (non-symbolic) numerical magnitudes (ANS-acuity) exhibits strong inter-personal variance (Halberda et al., 2008). In a very influential article, Halberda et al. (2008) reported that the inter-personal variance in ANS-acuity correlates with standardized mathematical performance tests in children. Halberda et al. (2008) speculated that the association between ANS-acuity and formal math achievement might be explained either by formal math education honing the ability to discriminate numerosities, or by the ANS being a “building block” for later formal mathematical knowledge. Especially the latter explanation has found a lot of attention (c.f. He et al., 2016); also reciprocal enhancement of both approaches are discussed (Elliott et al., 2019; Piazza et al., 2013). However, the association between ANS-acuity and formal math achievement is considered controversial, with some studies finding corresponding associations (Bonny & Lourenco, 2013; Libertus et al., 2011; Wang et al., 2017), while others did not (De Smedt et al., 2013; Sasanguie et al., 2014). A recent meta-analyses quantified an average effect for the association between non-symbolic number estimation tasks and mathematical competence measures with $r = .24$ (Schneider et al., 2017), i.e., pointing into the direction of an association with a small effect. However, this meta-analysis integrated results from a broad definition of mathematical competence and assessment methods, e.g., results from ANS-acuity assessments of various task paradigms and diverse implementations. Inconsistencies in the findings may be partially attributed to methodological diversity in the assessment of ANS-acuity, which is a critical factor, since not all assessment techniques have been found to be equally suitable (Dietrich et al., 2015).

2 Research Gap and Motivation for our Studies

A profound measurement of ANS-acuity might (partially) assess the potential of a high-level ability, i.e., mathematical competency, at the perceptual level. Due to the potential practical relevance regarding a relation between ANS-acuity and mathematical achievement, an understanding of the underlying mechanisms and assessment of the ANS ability proves crucial. The understanding of mechanisms likewise fosters better assessment practices and therefore increases validity. However, ANS mechanisms and theoretical assumptions are in need of deeper exploration, since important aspects are not empirically tested.

2.1 Assessment of ANS-Acuity

As previously mentioned, there is a wide range of assessment methods to cover ANS-acuity that vary in modality, paradigm or indices (dependent variables). A point of concern here is that studies with emphasis on methodological aspects found that different task variants and indices might not measure the same construct (Inglis & Gilmore, 2014; Price et al., 2012). If different ANS-tasks partially measure different constructs (even within the same modality), this poses a general concern for validity[2]. In our own studies (Ziegler et al., 2023; Ziegler & Drewing, 2022), we drew on empirical methodological work and theoretical considerations that discussed various methodological aspects of ANS-acuity measurement, which allowed an informed choice for our experiments (Dietrich et al., 2015; Inglis & Gilmore, 2014; Price et al., 2012).

Methods to assess a person's ability to discriminate numerical magnitudes follow different task paradigms and use different dependent variables (Dietrich et al., 2015). Tasks could potentially involve either direct numerosity estimation, representation(s), manipulation or comparison, since they are functions of the ANS (Odic & Starr, 2018). However, ANS-acuity assessment usually focuses on methods that avoid or minimize the involvement of other cognitive capacities (besides the numerosity processing) such as memory or the necessity of calculation. Frequently used methods to measure an individual's Approximate Number System acuity are dot comparison tasks (Dietrich et al., 2015). In these dot comparison tasks (DCTs), participants compare two sets of dot arrays (distinguishable by a feature or presentation mode) within a restricted time period and choose which array contained a higher number of dots. In general, DCTs follow one of three different paradigms: a paired presentation (massed presentation of number stimuli), a sequential presentation, or an intermixed presentation (Dietrich et al., 2015). In the paired presentation, the numerosity stimuli, e.g., dot-patterns, are presented simultaneously in different spatial locations, e.g., left and right side of a screen (massed). The sequential presentation uses a stream of number stimuli that are consecutively presented with a temporal delay, e.g., an audio sequence of tones. In the intermixed presentation mode, number stimuli are presented within the same spatial field, but can be distinguished by a feature that segregates the distinct entities. An example of an intermixed paradigm would be the presentation of blue and yellow dots within a shared visual field on a computer screen (see Halberda et al., 2008). Another relevant class of task paradigms are variants

---

[2] "Validity" here can be interpreted in a broader sense as compromising the theoretical assumptions of the ANS.

of same-different tasks, or match-to-sample variants in which non-symbolic number stimuli, e.g., a dot pattern, are visually presented and participants have to find a corresponding target numerosity from different alternatives (e.g., Fu et al., 2022; Sella et al., 2013). Likewise relevant are direct numerosity estimations tasks, in which participants directly perceive number stimuli and give an exact number of how many objects (e.g., dots) they saw (e.g., Lemaire & Lecacheur, 2007). This method already involves the ability to map the approximate representations linguistically onto number words (Odic & Starr, 2018).

The different ANS-acuity task paradigms can be implemented with dependent measures that are either based on reaction times or the number of correct trials (accuracy). Reaction times (rt) and accuracy can be used as a dependent measure, assuming that shorter rts, and more correct trials, reflect higher ANS-acuity, since more precise representations allow faster responses, or an overall higher number of correct trials reflect more precise representations (across different ratios) (De Smedt et al., 2013; Dietrich et al., 2015). However, reaction time based measures have been found to be less favorable than accuracy-based measures for measuring ANS-acuity (Dietrich et al., 2015; Inglis & Gilmore, 2014). There are also some (derivate) measures of reaction times and accuracy that are less recommended; among these are certain indices derived from regression analysis, e.g., "numerical ratio effects" (NRE) or "numerical distance effects (NDE) (c.f. Dietrich et al., 2015). Inglis and Gilmore (2014) suggested to completely abandon NRE indices as a measure for ANS-acuity, since they show poor reliability and convergent validity. Accuracy and weber fractions, however, have displayed reasonable convergent construct validity and acceptable reliability over different task paradigm implementations (Dietrich et al., 2015; Inglis & Gilmore, 2014; Lindskog et al., 2013; Price et al., 2012). Therefore, both measures are reasonable candidates to measure ANS-acuity. Guillaume and Van Rinsveld (2018) discussed that weber fractions require intricate calculations, but for most use cases do not provide additional information over simple accuracy measures. This is a strong argument for using accuracy measures, which is in line with the suggestion of Inglis and Gilmore (2014), who argued from a reliability perspective. However, Guillaume and Van Rinsveld (2018) pointed out that only specific weber-fractions allow modelling specific contributions of numerical and non-numerical dimensions in an approximate numerosity task.

Reviewing evidence from literature, we concluded that a paired DCT and match-to-sample tasks with accuracy measures or weber fractions (e.g., according to the method proposed by DeWind et al. (2015)) would be a reasonable methodology to investigate approximate numerosity estimation and test what mechanisms might underlie numerosity perception. To my knowledge, this method has not been applied or adapted to data of ANS-tasks in the haptic modality before. In our first study (Ziegler & Drewing, 2022), we went for a weber fraction algorithm similar to that proposed in DeWind et al. (2015), since we were interested in modeling potential contributing factors to the numerosity percept.

In our second study (Ziegler et al., 2023), we used an adapted variant of a match-to-sample task and accuracy (number of correct trials) as index to assess a person's ANS-acuity. Match-to-sample tasks formerly displayed validity (Dietrich et al., 2015) and with our adaptations we could additionally investigate spatial influences onto the numerosity percept in our experimental setup. To my knowledge, our approach of testing potential effects of spatial influences in a cross-modal (haptic to visual) setup with a paradigm using massed numerical stimuli has not been adapted before.

2.2 What Factors Shape a "Numerosity Percept"?

A second point that motivated this work stems from recent publications, which found that participants' response behavior in approximate numerosity tasks is influenced by other factors than just number, such as the convex hull, area, spatial or density related factors of a dot array (Bertamini et al., 2016; Clayton et al., 2015; Tomlinson et al., 2020). Area as a factor, e.g., the cumulative surface area of a dot array, was more likely to be relevant in younger children (Lourenco & Bonny, 2017; Tomlinson et al., 2020). Spatial related contributions, however, such as the convex hull, have repeatedly been found to be a significant factor in approximate numerosity tasks, also in adult samples (Clayton et al., 2015; Clayton & Gilmore, 2015; Gebuis & Reynvoet, 2011; Pekár & Kinder, 2020).

Non-numerical information, e.g., the spatial area covered by non-symbolic stimuli or their spatial arrangement, seems to affect the numerosity percept in unimodal approximate numerosity tasks (Clayton et al., 2015; DeWind et al., 2015; Gilmore et al., 2016; Hendryckx et al., 2021; Szucs et al., 2013; Tomlinson et al., 2020). DeWind et al. (2015) formalized that influence onto approximate numerosity perception into a taxonomy of number, size and spacing. Their method allows to statistically quantify these factors and calculate a weber fraction that controls for these influences (DeWind et al., 2015; DeWind & Brannon, 2016). However, this quantification has not yet been adapted to other modalities, which would clarify whether or not number solely accounts for the response behavior of participants in a DCT in another modality.

2.3 Modality x Method in ANS Assessment

A third major point that motivated our studies was that research on larger numerosities has mainly focused on the visual modality. Only few studies have utilized auditory stimulus material, e.g., tone sequences (e.g., Izard et al., 2009) or sequential dot and tone sequences in same-different or arithmetic tasks (Barth et al., 2003, 2005). Hence, the assumption of a system that extracts numerosity independently of the sensory input is rarely addressed in different sensory modalities. Furthermore, studies that have examined the numerosity perception with approximate numerosities in different modalities have often crossed the task paradigm with the modality, e.g., comparing auditory tone sequences (sequential paradigm) with visually presented dot arrays (massed presentation paradigm). Several studies reported performance differences between modalities when presentation format is crossed with modality (e.g., Anobile et al., 2018; Barth et al., 2003). In these cases, it is no longer possible to clearly attribute differences in performance (ANS-acuity) either to methodological variance or to actual differences based on the numerosity perception. To take this even further, studies that reasonably addressed methodological pitfalls in the ANS-acuity assessment and kept modality and presentation format constant show heterogeneous results in the performance between modalities, sometimes showing associations between modalities and sometimes not (Tokita et al., 2013; Tokita & Ishiguchi, 2016).

2.4 Summary Research Gap and State-of-the-Art

Even though the assumption of an amodal abstract numerosity percept is sound in theory, there are inconsistencies, ambiguities and straightforward unknowns regarding this part of the ANS theory. Unknowns remain because only a small proportion of studies actually probed approximate numerosity estimation in any other modality than the visual. A claim for amodality of numerosity by the ANS, however, should include empirical evidence from any modality in which approximate numerosity perception can be probed. I argue that this important aspect of the theory is volatile, since evidence in the auditory modality is rare and evidence from the haptic modality even scarcer.

Ambiguity stems from methodological issues in some studies, for example, when paradigms are crossed with modality, which makes the results difficult to interpret. Variance of the method is confounded with variance of the numerosity estimate. In addition, there is a plethora of methods in how to assess a person's ability to discriminate numerical magnitudes (ANS-acuity), but not all have been found to be reliable or valid measures. Further ambiguities come from studies that show heterogeneous results regarding the question whether the ANS enables a shared, perhaps abstract numerosity representation. Inconsistencies with the theoretical ANS assumptions come from recent studies in the visual modality that demonstrated that other factors than just "number" seem to be relevant contributors for a numerosity percept.

3 The Rationale of the Experiments

The overarching aim of this work was to test and clarify theoretical assumptions of the Approximate Number System. We critically investigated the amodality claim using symmetric methodology in behavioral experiments for the haptic and the visual modality. The haptic modality is suitable for experiments with approximate number, since task paradigms with massed presentation of numerical stimuli, like in the paired DCT or in a match-to-sample task, can be implemented. By including this modality and pairing it with state-of-the-art adapted methods, we enrich the current debate around approximate numerosity processing with new evidence from an underrepresented modality. Our chosen set of methods also allowed exploring whether spatial information in particular, e.g., certain dot pattern configurations, influence the response behavior of participants.

In the first study (Ziegler & Drewing, 2022), we focused on the association between performance in a DCT in the haptic and the visual modality with the idea of "convergent construct validity" in mind. If a numerosity percept is abstract and the source modality is irrelevant to the numerosity percept, as it is according to strongly formulated ANS assumptions, we expect a strong correlation across modalities between the performances of individuals in a DCT. In this work, we emphasized on using what we have elaborated as best practice in the assessment of ANS-acuity and use this knowledge to design a DCT with reasonably matched methodology in both modalities. As a novelty, we implemented a test-retest design for the haptic (as well as the visual) DCT, which allowed to evaluate the reliability of this implementation. Furthermore, our methods allowed for a differentiated analysis of the data, which revealed that indeed, and as expected, number was the primary factor affecting the response behavior of our participants. However, in both modalities, we also found that besides number another factor significantly influenced participants' performance: the spacing of the stimuli.

In our second study (Ziegler et al., 2023), we addressed more thoroughly the idea that humans extract numerosity from a certain source stimulus by "filtering out" covariant factors, such as spatial information, which was found to be significant in our first study (Ziegler & Drewing, 2022). According to the strong theoretical ANS assumptions, these spatial factors would not influence a participant's performance in another target modality. In this study, we used a novel form of numerosity matching task (an adapted match-to-sample task) in a cross-modal (haptic / visual) paradigm. We presented dot patterns haptically that conveyed spatial information allowing us to test whether spatial information does influence response behavior, contrary to the assumed mechanism in the strong assumptions of the ANS theory.

3.1 Summary of Study 1: Get in touch with numbers – an approximate number comparison task in the haptic modality[3]

Ziegler, M. C., & Drewing, K. (2022). Get in touch with numbers – an approximate number comparison task in the haptic modality. Attention, Perception, & Psychophysics. https://doi.org/10.3758/s13414-021-02427-6

In this study, we investigated the ANS amodality claim by means of a paired dot comparison task. We tested whether ANS-acuity measured with a paired DCT would yield similar results in the haptic and the visual modality; additionally, we tested the reliability of our task implementation. Furthermore, we were interested in potential factors that contribute to a numerosity percept other than number. Our screening of relevant literature identified two issues with the current state-of-the-art ANS assumption of amodality. First, studies that have investigated the perception of approximate numerosity using reasonable methodology found that responses of participants are influenced by other factors, such as spacing-related factors, and second, there is a lack of studies that tested the amodality assumption in another modality than the visual.

Especially the haptic modality has been underrepresented so far. Only a single study by Gimbert et al. (2016) used an adaption of a paired DCT with accuracy as performance index, which is well suited to investigate ANS-acuity. Gimbert et al. (2016) presented ratio-dependent performance as evidence for a shared cognitive system (ANS) that processed numerosity in both modalities, but they omitted an essential part of what we considered a conclusive empirical evidence for a shared (amodal) approximate number system: a strong correlation of their performance indices between modalities. We found that this and several other methodological aspects of the DCTs, especially their haptic DCT, could be improved. In brief, 1) we ensured sufficient spacing between dot stimuli to allow participants to reliably individuate the numerical stimuli with the palm of their hands. Second, we implemented a higher stimulus variety, i.e., different patterns, to reduce potential risk of biases. Third, we emphasized differentiated assessment and reporting of the tested ratios in the DCT that consecutively allowed a check for psychometric features (e.g., a reliability estimation). Lastly, we additionally calculated weber fractions adapted from DeWind et al. (2015) as ANS-acuity indicators as this index is a statistically adjusted weber fraction based on numerosity. It also allowed us to examine and model additional factors (e.g., the spatial information of a pattern) that might influence participants' responses.

In summary, with our implementation of a visual and a haptic DCT, we addressed the following questions in Ziegler and Drewing (2022):

a) Do adult participants express typical ratio-dependent performance in both versions (haptic, visual) of the DCTs?
b) What is the relative weight of factors that lead to the percept of numerosity?
c) How stable is the test-retest reliability of the haptic DCT and the visual DCT?
d) Do participants' individual and group-wise performances correlate between the haptic and the visual DCT (construct validity)?

---

[3] Transparency disclosure: Figures and Tables and their respective captions in this section are either directly taken from Ziegler and Drewing (2022) or slightly altered in their layout. License: CC BY 4.0. See references for source.

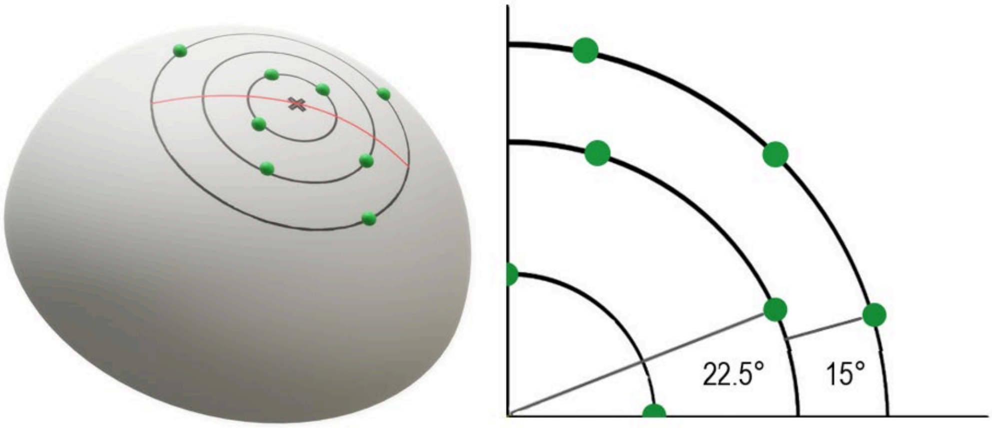
Methods

In the experiment, 67 participants performed two variations of a DCT, a visual and a haptic DCT, in each of two sessions. They had to determine which of the two dot arrays contained a higher number of dots. We tested different ratios in the DCT, 2.00, 1.33, 1.20, 1.14, and 1.11. The dot patterns contained 5 to 20 dots per stimulus. We referenced the task parameters on the implementation of Halberda et al. (2008).

A custom-made C++ program realized the visual DCT. We presented two dot patterns, one on the left side and one on the right side on a computer screen. The haptic DCT version was realized with styrofoam hemispheres and industrial map pins as numerical stimuli (see Figure 1). Participants sat at a table behind an opaque curtain, which prevented the participant from seeing the two hemispheres. Participants were instructed to examine the hemispheres with the palm of their hands.

Figure 1

Model of a Haptic Stimulus in the Experiment



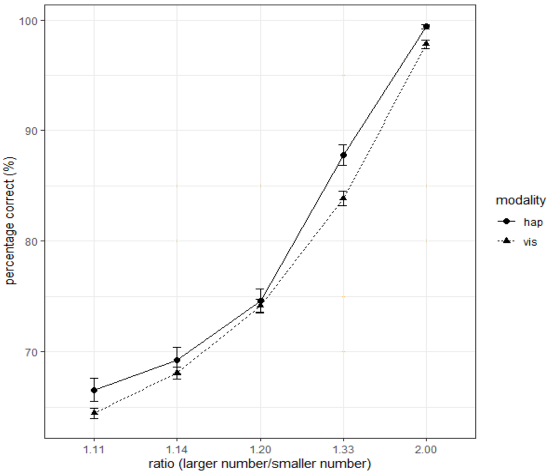
Note. Left Side: Styrofoam hemisphere with exemplary dots (pins) on three radial orbits. Right side: Quadrant with all possible slots for a pin. Position of the slots in degrees from the origin for one quadrant. Slots in the first orbit are perpendicular to the origin.

Results

Ratio dependency. We demonstrated that in both modalities, participants expressed a typical ratio-dependent response pattern. Smaller ratios became increasingly more difficult to distinguish for the participants. Furthermore, we found a significant main effect for modality (sensory modality), which reflected that participants showed better performance on average in the haptic modality (rmANOVA with the variables number ratio (2.00, 1.30, 1.20, 1.14, and 1.11) and sensory modality (haptic, visual), $F(4, 264) = 1367.56$, $p < .001$, $\eta_p^2 = .954$, all contrasts significant after Bonferroni adjustment, see Figure 2). The ratio dependence in both modalities can be seen as indicator for the involvement of the ANS.

Figure 2

Ratio-to-Correct Responses Plot for Visual and Haptic DCTs Aggregated Over Sessions



Note. Semilog plot. Errorbars represent the standard error of the mean.

Relative weight that contribute to the numerosity percept. We modeled our data for the visual and the haptic dataset according to adapted equations from DeWind et al. (2015) method. This allowed us to model the response behavior of all participants as a group model for the factors number, spacing and the tendency to choose a particular side. Table 1 shows the statistics of the underlying logistic regression model for the modalities (visual in two separate versions according to a control variable: DSC = dot size controlled / SAC = surface area controlled). The results show that besides number as strongest contributor the factor spacing became significant in each model, which indicates that spatial features (here mainly driven by the convex hull) are relevant for the numerosity percept.

Table 1

Estimated Model Coefficients and Statistics for Coefficient Tests Against 0 for Each Modality

Coefficient	β	Z	SE	p
Haptic				
β_{num}	2.534	48.39	0.052	<.001***
$\beta_{spacing}$	0.264	10.38	0.025	<.001***
β_{side}	-0.101	-6.82	0.015	<.001***
Visual DSC				
β_{num}	2.349	103.462	0.023	<.001***
$\beta_{spacing}$	0.267	41.917	0.006	<.001***
β_{side}	-0.010	-1.457	0.007	.145
Visual SAC				
β_{num}	2.045	101.537	0.020	<.001***
$\beta_{spacing}$	0.270	43.620	0.006	<.001***
β_{side}	-0.017	-2.568	0.007	.010[a]

Note. β_{num} and $\beta_{spacing}$ are regression weights for the log_2 number and spacing ratios, respectively, and β_{side} reflects a response bias towards one side; z is the test statistics of a significance test against zero, *SE* is the standard error of the regression weights. Visual condition 48,240 trials; haptic condition 10,720 trials. [a] not significant after Bonferroni adjustment.

Reliability and (Construct) Validity. Our check for the reliability of the DCTs revealed that the visual DCT showed acceptable reliability, while the haptic version suffered from relatively poor reliability (see Table 2), but was also measured with a smaller number of trials. We could not find correlations between the individual performance of the haptic and the visual DCTs of participants. Statistical corrections indicated that a correlational effect remained unlikely even under the assumption of higher reliability of the haptic DCT, based on the correction formulas. Hence, we found no indication for a construct validity that is in line with the classical assumptions of a strong ANS theory that would allow shared number processing between sensory modalities.

Table 2

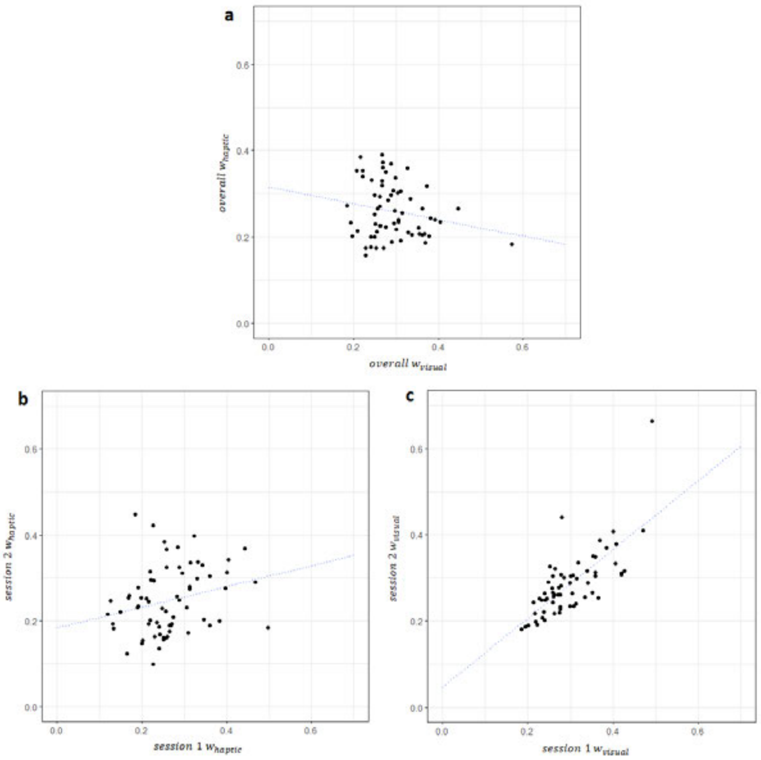
Correlations of ANS Acuity (w) Across Individuals, Test Statistics and Adjusted Task Reliability (ATR)

Modality	*r*	*N*	*t*	*df*	CI 95% [lower]	CI 95% [upper]	*P*	*ATR*
haptic – haptic	.251	67	2.088	65	0.011	0.463	.041*	.401
visual – visual	.736	64	8.549	62	0.598	0.831	<.001**	.848
haptic – visual	-.133	65	-1.067	63	-0.365	0.114	.290	-

Note. Intra-modal correlations are calculated between first and second session; cross-modal correlations are calculated between modalities (trials aggregated over both sessions, respectively). Spearman-Brown corrected reliability (for the number of trials in the cross-modal correlation) is reported. Statistics calculated for 80 trials.

Figure 3

Scatterplots of Individual w's (ANS-Acuity) for the Visual and the Haptic DCTs



Note. **a** visual-haptic, **b** haptic t1-t2, and **c** visual t1-t2. Intra-modal correlations are calculated between first and second session; cross-modal correlations are calculated between modalities (aggregated over both sessions; t1 = session 1, t2 = session 2, respectively).

Discussion

In this work, we demonstrated that regardless of the source modality (haptic or vision), participants mainly based their responses on number in the DCTs, which is indicated by the ratio-dependent performance, as well as the regression weights in our logistic regression model. However, the influence of number slightly but significantly differed between the sensory modalities and, additionally, we found for both modalities a similarly sized smaller influence of spacing-related factors (driven by convex hull). The reliability of our DCTs turned out to be reasonable in the visual modality but relatively poor in haptic modality. Our analysis did not reveal a correlational link between the performances of the two modalities even though number was a strong influence in all conditions. This absence of an effect remained after accounting for poor haptic task reliability. In a strong interpretation of the ANS theory, number is an independent feature and the ANS would filter out any non-numerical stimulus features from the source material. Therefore, the performance of participants should have been highly similar between modalities, especially since we used the same numerical content in the same assessment paradigm. However, our results did not indicate an association of ANS-acuity between the modalities. With the methodological improvements we made in design, haptic stimulus material, and particularly in data analysis, we conclude that the assumption of a common cognitive system for numerosity encoding between the visual and haptic modality might have been made precipitously. Our results might be explained in a theoretical framework that allows other features than just number to shape a numerosity percept and include modality-specific mechanisms.

3.2 Summary of Study 2: The Role of Spatial Information in an Approximate Cross-Modal Number Matching Task[4]

Ziegler, M. C., Stricker, L.K., & Drewing, K. (2023). The role of spatial information in an approximate cross-modal number matching task. Attention, Perception, & Psychophysics. https://doi.org/10.3758/s13414-023-02658-9

In our first study, we found rather contradictory evidence to the widely assumed axiom of amodality that goes along with “strong” interpretations of the ANS theory. We confirmed that numerosity processing in the haptic modality is possible and, furthermore, that spatial information (factor spacing) was a significant predictor as well. In this study, we elaborated on that finding in using a cross-modal paradigm to clarify whether spatial information is extracted from non-symbolic numerical stimuli and influences the discrimination performance of participants. Previous cross-modal studies often times relied on sequential presentation of numerical magnitudes, e.g., via vibro-tactile stimuli that present numerosity with sequences of vibrations (Tokita & Ishiguchi, 2016), which, unlike massed presentation, e.g., in a dot-array, does not convey topological, spatial information. Here, we adapted a number match-to-sample task that allowed to include spatial information and consequently test whether it is relevant. Our idea was that if spatial information is extracted and affects individual’s responses, it contradicts the widely assumed ANS mechanism. Specifically, we tested whether numerical and spatial information was extracted from the haptic modality and whether the information is available to participants when making a numerosity comparison in the visual modality.

---

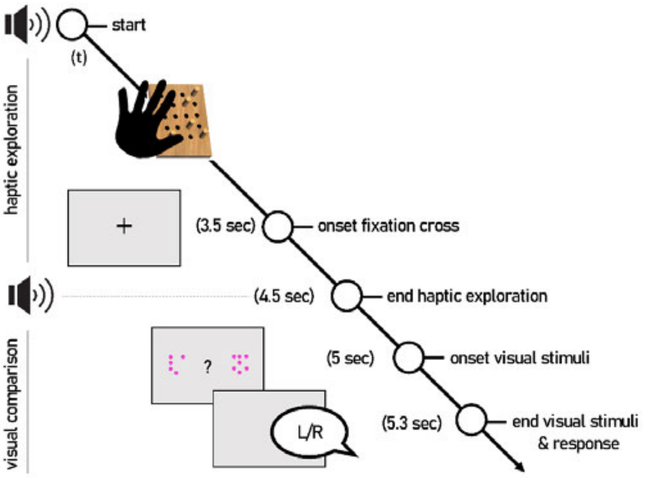
[4] Transparency disclosure:
This paper partially incorporates data that also has been used for L.K. Stricker's bachelor's thesis (Stricker, 2021), which was supervised by MCZ.
Figures and Tables and their respective captions in this section are either directly taken from Ziegler et al. (2023) or slightly altered in their layout. License: CC BY 4.0. See references for source.

Methods

50 undergraduates participated in our cross-modal number matching task. Participants extracted numerosity from a dowel array that was explored with the palm of the hand for a total of 4.5 seconds. Afterwards, they compared the numerosity to two visually presented dot arrays (see Figure 4). Participants were instructed to decide which of the two visually presented patterns matched the numerosity that was prior presented in the haptic modality. We varied whether the visual target stimulus was a random dot array (spatially random [SR]) or the exact spatial layout of the haptic dowels (spatially identical [SI]). In the spatially random condition, we furthermore manipulated the convex hull of the dot patterns and evaluated whether convex hull (more specific, an index based convex hull and sparsity of the array) influences the participants responses. Additionally, we varied the difficulty of the pattern match by varying the ratio between the target stimulus and the distractor stimulus (1.11, 1.14, 1.20, 1.33. 2.00).

Figure 4

Illustration of the Procedure in the Cross-Modal Matching Task.



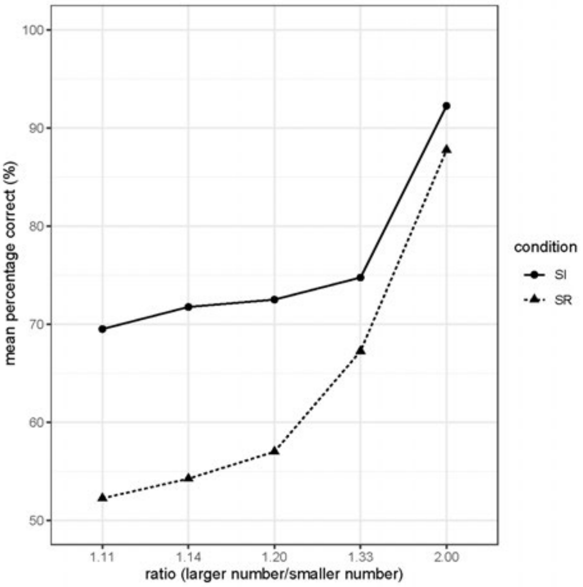
Results

Ratio and Pattern. We ran a rmANOVA that revealed a significant main effect for the factor ratio (2.00, 1.30, 1.20, 1.14, and 1.11) with $F(4,196) = 59.246$, $p < .001$, $\eta_p^2 = .547$ and a significant main effect for the factor pattern (spatially identical, spatially randomized) with $F(1,49) = 59.571$, $p < .001$, $\eta_p^2 = .549$. In general, participants scored with higher accuracy with higher ratios, which is typical for numerosity processing. However, the performance was significantly elevated for spatially identical patterns (see Figure 5). Furthermore, there is a significant interaction effect between ratio and pattern, $F(1,196) = 2.791$, $p < .001$, with a smaller effect, $\eta_p^2 = .054$.

Logistic Regression Analyses. Within the subset of trials in the SR pattern condition, we distinguished trials where the haptic stimulus was spatially congruent vs. spatially incongruent to the visual target stimulus. We calculated two logistic regression models, each for trials in which the target number was higher (model A) than the distractor and complementarily for the lower target number (B). We found that for model A, spatial congruency (based on convex hull) facilitated correct responses whereas in model B, spatial congruency was not a significant predictor. Only high numerical magnitude, paired with high congruent spatial magnitude, resulted in facilitating the performance of participants.

Figure 5

Mean Percentage Correct per Ratio and Pattern in the Complete Sample.



Note. "SI" refers to the spatially identical pattern condition when the haptic dowel array and the correct visual dot array match in numerosity and also in spatial arrangement. "SR" refers to the spatially randomized pattern condition, when the haptic dowel array and the correct visual dot array solely match in numerosity but not in spatial arrangement. *N* = 50.

Discussion

In this study, we again confirmed that individuals are able to use numerical information extracted from haptically presented source material. In addition, we demonstrated that participants are able to match the haptically extracted numerosity to a visually presented target with typical ratio-dependent accuracy that is related to the difficulty of the numerosity comparison in a cross-modal paradigm. Since we used a paradigm with massed numerical stimuli, we were able to investigate and manipulate spatial properties of the patterns. As an improvement to our previous study, we varied the haptic (and also the visual) patterns in every trial, which has been discussed in Ziegler and Drewing (2022). In a "strong" interpretation of abstract numerosity estimation within the classical ANS theory assumption, spatial information would not affect performance under manipulation of the spatial layout of the arrays. However, contrary to this assumption, our analyses revealed a significant main effect of the factor "pattern", which implies that spatial information is a) extracted b) maintained throughout cognitive processing and c) impacts a participant's response behavior. This result again encouraged us to think and discuss alternative proposals that would account better for additional non-numerical information to influence participants' responses. We also discussed whether attentional factors might contribute to our results and whether there is a more general framework in which approximate number processing can be located: ensemble statistics (c.f. Alvarez, 2011).

4 General Discussion

The Approximate Number System is thought to be part of the cognitive system which allows to perceive and represent number (Brannon & Merritt, 2011; Dehaene, 2011; Feigenson et al., 2004; Odic & Starr, 2018; Spelke, 2022; Spelke & Kinzler, 2007). One assumption of the ANS is that humans can form abstract representations of number regardless of the source modality, e.g., extract the (approximate) number of coins in the palm of a hand, both visually or haptically (Brannon & Merritt, 2011). According to the most prevalent theory, the ANS extracts numerosity directly from the environment by a mechanism that filters out covariant features (Dehaene, 2011; Gebuis et al., 2016; Leibovich, Katzin, Harel, et al., 2017; Lourenco & Aulet, 2022; Tokita & Ishiguchi, 2016). This aspect of numerosity extraction of the ANS theory can be classified as "direct model" of numerosity estimation, in which "number" is a primary feature of a stimulus, rather than a feature derived from auxiliary cues, e.g., density or area, which in turn can be categorized as "indirect models" of numerosity perception (Qu et al., 2021). Currently, the yet unclear underlying mechanisms of numerosity perception are "hotly" debated (Testolin et al., 2020).

In our studies, Ziegler and Drewing (2022) and Ziegler et al. (2023), we tested and clarified theoretical assumptions of the ANS with an emphasis on using different modalities (haptic / visual) and elaborated methods for ANS assessment. In particular, we investigated the amodality claim of the ANS and the assumed underlying mechanism that extracts numerosity directly from the stimuli regardless of the source modality into an abstract number percept. We adapted methodology, which has already displayed validity in capturing individual differences in numerosity processing (Dietrich et al., 2015; Inglis & Gilmore, 2014; Price et al., 2012). In addition, we kept the presentation format for the number stimuli constant instead of crossing the presentation format with modality, which confounds potential method and modality effects.

Since studies that investigated numerosity perception in the haptic modality are very sparse, one important result of our studies is that approximate numerosity processing is possible with means of paired or massed presentation of approximate numerical stimuli in the haptic modality. In both of our methodologically adapted paradigms, intra-modal and cross-modal with massed presentation of numerical stimuli, we found typical ratio-dependent performances of participants, which is typical for processing of number (Brannon & Merritt, 2011; Dehaene, 2011; Feigenson et al., 2004; Gimbert et al., 2016; Odic & Starr, 2018). Ratio-dependent performance in a numerosity comparison task is a strong indicator of ANS involvement and has been shown intra-modally before in the combination of the visual and haptic modality with massed presentation of numerical stimuli (Gimbert et al., 2016).

Yet, our results in Ziegler and Drewing (2022) and Ziegler et al. (2023) indicate that participants mainly, but not exclusively, used numerosity to accomplish the respective task. We demonstrated additional influence of spatial factors with a logistic regression model adapted from DeWind et al. (2015) in Study 1, and with an experimental paradigm that systematically varied spatial information in the number-matching task in Study 2.

Assuming that number is the primary feature and the ANS filters out all confounding features from the stimulus material, an individual's ANS-acuity should have been highly similar across modalities. We investigated this in Study 1 as an instance of (construct) validity for a shared amodal system. Our results in Ziegler and Drewing (2022), however, did not reveal a correlation of performance (weber fractions) between modalities, which would have been plausible if participants discriminated abstract representations of approximate numerical magnitude within the task. The individual's weber fractions rather differed, not only in the haptic condition compared to the visual condition but also intra-modally within the visual control conditions. Some differences may be partially attributed to modality specific

DCT implementations we made, but since we have reasonably adapted the tasks in our experiment, we think performance differences rather reflect actual differences in numerosity processing. However, there is a caveat when interpreting the association of weber fractions between the haptic and visual modality in our first study. We investigated the test-retest reliability of both paired DCTs (visual / haptic), which revealed sufficient visual DCT reliability but a rather low haptic DCT reliability (c.f. Evers, 2001). This is a methodological shortcoming, which may be attributed to the number of experimental trials in the experiment, which was lower for the haptic DCT. Therefore, we could not unequivocally draw a conclusion about an association of ANS-acuity between the modalities. However, even after a correction for attenuation we did not find a medium-sized or larger (positive) correlation between the weber fractions (Ziegler & Drewing, 2022). Differences in the numerosity perception between modalities and between visual conditions furthermore became visible (and tangible) when we modeled the responses with the prediction model of DeWind et al. (2015). The probability to choose a certain side as a function of numerosity (number ratio) differed in the haptic and the visual modality modeled over all participants. This evidence also questions the independence of the numerosity percept from the source material and rather points towards modality specific mechanisms (Ziegler & Drewing, 2022). At last, in Study 1, we found evidence that factors beyond “number ratio” (i.e., numerosity) contributed to the participants’ responses. In particular, we found that the compound factor of “spacing”, which reflects spatial information of a pattern, was a significant predictor in our DCTs. In our second study, we further explored the effect of spatial information onto the participants’ performance, when we directly manipulated the spatial configuration of dot patterns (or dowel patterns). The results showed a striking facilitating effect for patterns that were identical in their layout between the haptic and the visual modality. Furthermore, we found a (partial) facilitating effect when non-identical patterns were spatially congruent (Ziegler et al., 2023). These facilitation effects are remarkable, since they indicate that spatial information was transferred via the haptic source modality and affected the response behavior of participants even across modalities. Spatial information of the numerical source stimuli was apparently extracted along with number, which is another dissonance to the classical understanding of ANS mechanisms as described in e.g., Brannon and Merritt (2011); Dehaene (2011); Dehaene and Changeux (1993).

Taking evidence of both of our studies in account, we consider the assumption of a mechanism that enables abstract perception of numerosity by filtering out covariant stimulus features to be unlikely. Our studies demonstrated empirical evidence that allows the interpretation that a “numerosity percept” is sensitive to stimulus features that can be modality-specific but also occur within the same modality depending on the controlled stimulus feature (e.g., average dot size vs. cumulative surface area). I speculate that these subtle differences in the numerical stimuli shape numerosity percepts, which diverge to the extent that we could not find an association of numerosity discrimination performance between the haptic and visual modalities (Ziegler & Drewing, 2022). However, they do not discount that number is the strongest influencing factor, but necessarily question the “cardinality” of the mechanism of approximate numerosity perception proposed by the “strong” assumption of the ANS theory.

The assessment of ANS-acuity, or in more general, the ability to discriminate approximate numerical magnitudes, with reasonable methodology has been a focus in our studies. Our methodological adaptation of paradigms that have been found valid in the visual modality (Dietrich et al., 2015) allowed to test an assumed axiom of the ANS theory. Vice versa, our experiments revealed mechanisms that contribute to a numerosity percept in both the visual and the haptic modality and across both modalities. Our findings suggest paying even more attention to the implementations of the numerosity tasks, since performance appears to be affected by various stimuli features. This is in line with recent meta-analytical findings of Guillaume and Van Rinsveld (2018), who expressed concerns that different control conditions significantly affected performance in numerosity tasks, weber fractions in particular. Results of both our studies showed that the non-numerical dimensions affected performance, most likely in both of our examined modalities.

However, the modality-specific adaptations of massed presentation of “dot” arrays also have their own methodological limitations[5], which I want to point out here, since our haptic implementations are a novelty. In our haptic numerosity tasks, we focused on important details, e.g., assuring that the spacing of stimuli (pins or dowels) allows participants to enumerate every single entity with the palm of their hands. We ensured this via pre-tests and screening the scarce suitable literature (c.f. Craig & Lyle, 2001; Ginsburg & Pringle, 1988). However, we did not implement a control condition equivalent to the cumulative surface area control (SAC-control) implemented in some visual numerosity estimation tasks (c.f. Halberda et al., 2008) and also in a haptic numerosity discrimination task of Gimbert et al. (2016). The main issue we saw here, which ultimately led to the decision not to implement a direct equivalent as in the visual domain, is that the haptic skin deformation evokes the “intensity” in active touch (c.f. Hayward, 2011). Consequently, smaller pins would rather cause more intense stimulation than larger ones, following an inverse logic of the surface-area control in the visual modality (Ziegler & Drewing, 2022). For future research, however, this is an interesting and important topic to cover, i.e., does intensity of skin deformation of different numerical stimuli systematically modify a numerosity percept? Overall, I think that the haptic modality is an adventurous testing ground for numerous approximate numerosity related paradigms, e.g., paired paradigms like in our studies. I think it is exciting to endeavor the modality further as it can also be utilized to test whether numerical illusions or adaptation effects that have been found in the visual modality also occur haptically or cross-modally (c.f. Adriano et al., 2021).

Another general criticism of both of our studies is that participants may have used “strategies” to some extent in the respective task which “bypassed” the mechanisms of numerosity estimation. Such strategies would be rather domain unspecific and might have contributed to the task performance. Evidence from literature indicates that there is indeed inter-individual use of strategies in numerosity estimation tasks by participants (Dietrich et al., 2019). In our second study (Ziegler et al., 2023), for example, the facilitating effect found in the “spatially identical” condition could partially stem from a strategy use of “shape recognition”, which can be seen as a factor that bypassed approximate numerosity estimation. In other words, participants would have achieved better discrimination performance because they simply identified salient points of the dowel pattern and matched the “shape” of the pattern to the visual target, rather than estimating the numerosity. However, considering the bigger picture of evidence (ratio dependency, logistic regression analyses in both studies; Ziegler et al., 2023; Ziegler & Drewing, 2022), I think that participants mainly processed numerical and spatial attributes of the dot stimuli to meet task demands, and that this information influenced internal representation of numerosity rather than extensive strategy use. Since “shape recognition” is also rarely reported as a strategy in numerosity estimation tasks (Roquet & Lemaire, 2019), I think of strategy use as possible but rather minor factor in our experiments. Further studies, however, should investigate whether the frequency of the described strategies found in the visual modality applies similarly to the haptic modality, and qualitatively examine whether certain strategies are used to estimate numerosity (see Dietrich et al., 2019; Roquet & Lemaire, 2019).

---

[5] Specific limitations can be found in the respective papers. I focus here on aspects that can apply for both studies.

As already mentioned, the underlying mechanisms of (approximate) number perception and processing are currently "hotly" debated (Testolin et al., 2020) and different proposals have been made, which are likewise controversially debated. These proposals add perspectives on numerosity perception but differ in their views whether number is a distinct, primary feature, abstract, a magnitude among others or a feature that is composed from non-numerical magnitudes of stimuli or whether there is an innate "number sense" (Allïk & Tuulmets, 1991; Clarke & Beck, 2021; Dehaene, 2011; Gebuis et al., 2016; Gevers et al., 2016; Kadosh & Walsh, 2009; Leibovich, Katzin, Harel, et al., 2017; Lourenco & Aulet, 2022; Walsh, 2003; Zorzi & Testolin, 2018). Our research contributes to the forthcoming in the field since we revealed mechanisms that shape a numerosity percept in the haptic and the visual modality with coherent and plausible methodology. Furthermore, we showed numerosity percepts can deviate from modality to modality and within control conditions. Our results are in line with recent findings that revealed additional influences in primarily visual numerosity tasks (c.f. Clayton et al., 2015; Clayton & Gilmore, 2015; Gilmore et al., 2016; Tomlinson et al., 2020). In both our studies, we accumulated evidence that shows in particular that spatial and potentially other factors, e.g., area, influence a percept. As mentioned, we think that our results are partly in conflict with the classical ANS view and that other theoretical proposals could generally explain our data quite well. For example, the "sensory integration" proposal given by Gebuis et al. (2016) would in principle be in line with our data. Their proposal can be framed as an "indirect model" of numerosity estimation and suggest that numerosity estimation is a process of integrating different sensory cues, e.g., density of a dot pattern or its surface area. According to their view, different sensory cues contribute to a final percept with a specific weight. Therefore, features that are more salient would gain more significance in deciding which dot array is more numerous. This proposal is widely congruent with our observations and the results of our regression models, since Gebuis et al. (2016) also emphasize that sensory cues are not removed in the process of shaping a numerosity percept (opposed to the classical ANS view). Another important factor in sensory integration theory is that Gebuis et al. (2016) do not argue for "abstract" numerosity, in a sense of it being abstract when a mental representation is solely influenced by the numerical value. This could possibly explain the absence of a correlation between modalities we found in Ziegler and Drewing (2022). My speculation here is that the numerosity percept was weighted differently in both modalities due to "sensory cues", which resulted in rather modality-specific weights and representations of numerosity. If any "confounding" variable had influence on the percept and numerosity was estimated with an integration of the stimulus features, it would not be likely that our adapted DCTs would result in converging percepts (c.f. Ziegler & Drewing, 2022). Considering number itself can be perceived differently between modalities, a lack of performance correspondence between the DCTs in our first study would be comprehensible. To investigate the weighting of sensory cues, especially for the haptic modality, will be a challenge for further research. Modifications of our experimental setup in Study 2 could provide a first approach here. For example, one could use the method of adjustment to systematically vary a sensory feature along a single dimension, e.g., the convex hull of a pattern conveyed by a haptically presented dowel pattern to match a given numerosity, e.g., visually presented. These experiments would be intricate because each step of variation in the sensory feature would also need to be accompanied by another combination of the numerical stimuli that retains the same numerosity but varies in the sensory feature.

As formerly mentioned, the sensory integration theory is only one framework that could explain our data. Other authors also argue against the classical ANS assumptions and discourage the classical "number sense" theory, but still argue for number as feature that can "holistically" shape a "magnitude" percept along with other magnitudes

(Leibovich, Katzin, Harel, et al., 2017; Leibovich, Katzin, Salti, et al., 2017). This can be seen as an extension of the classical ANS perspective. Our data might also partially comply with this perspective, since the mechanisms of how a numerosity percept is derived are open to number, which contributed strongly to the percept in our experiments, as well as other magnitudes. Lourenco and Aulet (2022) have recently introduced an interesting perspective, which might also cover a plausible explanation for some of our findings. They argued that number and accompanying features of the non-symbolic numerical stimuli, such as spacing or area, are not filtered out in a "sensory stage" (see 1.1.2 Approximate Number System and its Hallmarks) but rather maintained as "magnitude" information encoded along with number. All information would therefore be conveyed and available to an individual. According to their proposal, the "experience" of number can be evoked by attention to an internal integral representation. This could, for example, explain the strong advantage for identical patterns we found in Ziegler et al. (2023), since all numerical and non-numerical information was available in a compound percept and covariant congruent information (spatial features of a pattern) facilitated the response. However, with this argument, the theory is already taken further than proposed by the authors, since Lourenco and Aulet (2022) explicitly limited their ideas to visual number processing.

The field of numerical cognition is not short on alternative proposals that try to explain aspects and phenomena of numerical processing and perception. We suggest there is accumulating evidence to shift the perspective from a system that purely and abstractly represents approximate numerical magnitudes into an updated version that acknowledges recent findings, e.g., DeWind et al. (2015); Gevers et al. (2016); Gilmore et al. (2016); Hendryckx et al. (2021); Leibovich, Katzin, Harel, et al. (2017); Smets et al. (2015). The debate about the cognitive and neural representations of number and the search for mechanisms in approximate numerosity perception will be an ongoing topic, with new perspectives emerging frequently (e.g., Clarke & Beck, 2021; Lourenco & Aulet, 2022; Pickering et al., 2023; Sixtus et al., 2023). I think the sheer number of published proposals and the intensity of the debate about one of the core mechanisms of numerical cognition are proof that the classical ANS theory (e.g., classical number sense) has served the field well, but that it must engage with new ideas and findings. From the data we gathered in Ziegler & Drewing (2022) and Ziegler et al. (2023), we will not argue for a specific proposal, since this is beyond the scope of our experiments. Our scope in both studies was on approximate numerosity processing in the visual and haptic modality, with methods of massed presentation of dot stimuli and the "strong" assumptions of the ANS theory. This covers and contributes to a small but substantial part in the overall discussion. Since our data could be better explained by other theories, we argue for a broader interpretation of the ANS that allows for additional factors other than number to be involved in shaping a numerosity percept or for a shift to alternative proposals (Ziegler et al., 2023). I think we have obtained results that advance the field and provide arguments for stimulus and modality-specific numerosity perception and against classical "strong" ANS theory assumptions, but further research is needed. We also encourage the idea proposed by Marinova et al. (2021), who argue for a specific theoretical framework that explains all diverse phenomena in numerical cognition, and consider that humans are rather competent of using multiple ways to estimate or process numerosity, depending on the individual and the context of the experiment (e.g., specifics of the modality). In addition to that outlook, as we discussed in Ziegler et al. (2023), I think that insights from approximate number perception could be reflected in the context of ensemble perception or ensemble statistics and attention research (c.f. Alvarez, 2011). Findings from this field suggest that the cognitive system pools local features from (visual) scenes into a statistical summary (which can be numerical and spatial) to handle attentional capacity limits (Alvarez & Oliva, 2009). In this context, it would be an interesting future topic to examine whether there are common mechanisms underlying numerosity representation and summary statistics (Alvarez, 2011), whether they may be concurrent mechanisms, and whether spatial and location-based attentional resources need to be better taken into account when designing approximate numerosity tasks, which indeed seems relevant (c.f. Pomè et al., 2021).

Taken together, even though the current state of knowledge on approximate numerosity processing has not yet revealed what factors underlie the (approximate) numerosity percept and what is therefore known as "ANS-acuity", the potential usefulness of measuring this ability remains. We strongly argue for common standards when assessing ANS-acuity, which include paradigms, indices, and implementation details (accurate stimulus protocol, e.g., control for convex hull or spacing) and, furthermore, considerations of a) modality specific effects and b) possible "carry-over" effects of stimulus features in cross-modal ANS tasks.

5 References

Adriano, A., Rinaldi, L., & Girelli, L. (2021). Visual illusions as a tool to hijack numerical perception: Disentangling nonsymbolic number from its continuous visual properties. *Journal of Experimental Psychology: Human Perception and Performance*, *47*(3), 423–441. https://doi.org/10.1037/xhp0000844

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

Allïk, J., & Tuulmets, T. (1991). Occupancy model of perceived numerosity. *Perception & Psychophysics*, *49*(4), 303–314. https://doi.org/10.3758/BF03205986

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, *15*(3), 122–131. https://doi.org/10.1016/j.tics.2011.01.003

Alvarez, G. A., & Oliva, A. (2009). Spatial ensemble statistics are efficient codes that can be represented with reduced attention. *Proceedings of the National Academy of Sciences*, *106*(18), 7345–7350. https://doi.org/10.1073/pnas.0808981106

Anobile, G., Arrighi, R., Castaldi, E., Grassi, E., Pedonese, L., Moscoso, P. A. M., & Burr, D. C. (2018). Spatial but not temporal numerosity thresholds correlate with formal math skills in children. *Developmental Psychology*, *54*(3), 458–473. https://doi.org/10.1037/dev0000448

Barner, D., & Baron, A. S. (2016). *Core Knowledge and Conceptual Change*. Oxford University Press.

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition*, *86*(3), 201–221. https://doi.org/10.1016/S0010-0277(02)00178-6

Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences*, *102*(39), 14116–14121. https://doi.org/10.1073/pnas.0505512102

Bertamini, M., Zito, M., Scott-Samuel, N. E., & Hulleman, J. (2016). Spatial clustering and its effect on perceived clustering, numerosity, and dispersion. *Attention, Perception, & Psychophysics*, *78*(5), 1460–1471. https://doi.org/10.3758/s13414-016-1100-0

Bisazza, A., & Gatto, E. (2021). Continuous versus discrete quantity discrimination in dune snail (Mollusca: Gastropoda) seeking thermal refuges. *Scientific Reports*, *11*(1), 3757. https://doi.org/10.1038/s41598-021-82249-6

Bonny, J. W., & Lourenco, S. F. (2013). The approximate number system and its relation to early math achievement: Evidence from the preschool years. *Journal of Experimental Child Psychology*, *114*(3), 375–388. https://doi.org/10.1016/j.jecp.2012.09.015

Brannon, E. M., & Merritt, D. J. (2011). Chapter 14—Evolutionary Foundations of the Approximate Number System. In S. Dehaene & E. M. Brannon (Eds.), *Space, Time and Number in the Brain* (pp. 207–224). Academic Press. https://doi.org/10.1016/B978-0-12-385948-8.00014-1

Cantlon, J. F., Cordes, S., Libertus, M. E., & Brannon, E. M. (2009). Numerical abstraction: It ain't broke. *Behavioral and Brain Sciences*, *32*(3–4), 331–332. https://doi.org/10.1017/S0140525X09990513

Clarke, S., & Beck, J. (2021). The number sense represents (rational) numbers. *Behavioral and Brain Sciences*, *44*. https://doi.org/10.1017/S0140525X21000571

Clayton, S., & Gilmore, C. (2015). Inhibition in dot comparison tasks. *ZDM*, *47*(5), 759–770. https://doi.org/10.1007/s11858-014-0655-2

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

Craig, J. C., & Lyle, K. B. (2001). A comparison of tactile spatial sensitivity on the palm and fingerpad. *Perception & Psychophysics*, *63*(2), 337–347. https://doi.org/10.3758/BF03194474

De Smedt, B., Noël, M.-P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, *2*(2), 48–55. https://doi.org/10.1016/j.tine.2013.06.001

Dehaene, S. (2011). *The number sense: How the mind creates mathematics* (Rev. and updated ed). Oxford University Press.

Dehaene, S., & Changeux, J.-P. (1993). Development of Elementary Numerical Abilities: A Neuronal Model. *Journal of Cognitive Neuroscience*, *5*(4), 390–407. https://doi.org/10.1162/jocn.1993.5.4.390

DeWind, N. K., Adams, G. K., Platt, M. L., & Brannon, E. M. (2015). Modeling the approximate number system to quantify the contribution of visual stimulus features. *Cognition*, *142*, 247–265. https://doi.org/10.1016/j.cognition.2015.05.016

DeWind, N. K., & Brannon, E. M. (2016). Significant Inter-Test Reliability across Approximate Number System Assessments. *Frontiers in Psychology*, *7*. https://www.frontiersin.org/article/10.3389/fpsyg.2016.00310

Dietrich, J. F., Huber, S., & Nuerk, H.-C. (2015). Methodological aspects to be considered when measuring the approximate number system (ANS) – a research review. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00295

Dietrich, J. F., Nuerk, H.-C., Klein, E., Moeller, K., & Huber, S. (2019). Set size influences the relationship between ANS acuity and math performance: A result of different strategies? *Psychological Research*, *83*(3), 590–612. https://doi.org/10.1007/s00426-017-0907-1

Elliott, L., Feigenson, L., Halberda, J., & Libertus, M. E. (2019). Bidirectional, Longitudinal Associations Between Math Ability and Approximate Number System Precision in Childhood. *Journal of Cognition and Development*, *20*(1), 56–74. https://doi.org/10.1080/15248372.2018.1551218

Evers, A. (2001). Improving Test Quality in the Netherlands: Results of 18 Years of Test Ratings. *International Journal of Testing*, *1*(2), 137–153. https://doi.org/10.1207/S15327574IJT0102_3

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Fu, W., Dolfi, S., Decarli, G., Spironelli, C., & Zorzi, M. (2022). Electrophysiological Signatures of Numerosity Encoding in a Delayed Match-to-Sample Task. *Frontiers in Human Neuroscience*, *15*. https://www.frontiersin.org/articles/10.3389/fnhum.2021.750582

Gebuis, T., Cohen Kadosh, R., & Gevers, W. (2016). Sensory-integration system rather than approximate number system underlies numerosity processing: A critical review. *Acta Psychologica*, *171*, 17–35. https://doi.org/10.1016/j.actpsy.2016.09.003

Gebuis, T., & Reynvoet, B. (2011). Generating nonsymbolic number stimuli. *Behavior Research Methods*, *43*(4), 981–986. https://doi.org/10.3758/s13428-011-0097-5

Gevers, W., Kadosh, R. C., & Gebuis, T. (2016). Chapter 18 - Sensory Integration Theory: An Alternative to the Approximate Number System. In A. Henik (Ed.), *Continuous Issues in Numerical Cognition* (pp. 405–418). Academic Press. https://doi.org/10.1016/B978-0-12-801637-4.00018-4

Gilmore, C., Cragg, L., Hogan, G., & Inglis, M. (2016). Congruency effects in dot comparison tasks: Convex hull is more important than dot area. *Journal of Cognitive Psychology*, *28*(8), 923–931. https://doi.org/10.1080/20445911.2016.1221828

Gimbert, F., Gentaz, E., Camos, V., & Mazens, K. (2016). Children's Approximate Number System in Haptic Modality. *Perception*, *45*(1–2), 44–55. https://doi.org/10.1177/0301006615614448

Ginsburg, N., & Pringle, L. (1988). Haptic Numerosity Perception: Effect of Item Arrangement. *The American Journal of Psychology*, *101*(1), 131–133. https://doi.org/10.2307/1422798

Gómez-Laplaza, L. M., & Gerlai, R. (2011). Can angelfish (Pterophyllum scalare) count? Discrimination between different shoal sizes follows Weber's law. *Animal Cognition*, *14*(1), 1–9. https://doi.org/10.1007/s10071-010-0337-6

Guillaume, M., & Van Rinsveld, A. (2018). Comparing Numerical Comparison Tasks: A Meta-Analysis of the Variability of the Weber Fraction Relative to the Generation Algorithm. *Frontiers in Psychology*, *9*, 1694. https://doi.org/10.3389/fpsyg.2018.01694

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

Hanus, D., & Call, J. (2007). Discrete quantity judgments in the great apes (Pan paniscus, Pan troglodytes, Gorilla gorilla, Pongo pygmaeus): The effect of presenting whole sets versus item-by-item. *Journal of Comparative Psychology*, *121*, 241–249. https://doi.org/10.1037/0735-7036.121.3.241

Hayward, V. (2011). Is there a 'plenhaptic' function? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1581), 3115–3122. https://doi.org/10.1098/rstb.2011.0150

He, Y., Zhou, X., Shi, D., Song, H., Zhang, H., & Shi, J. (2016). New Evidence on Causal Relationship between Approximate Number System (ANS) Acuity and Arithmetic Ability in Elementary-School Students: A Longitudinal Cross-Lagged Analysis. *Frontiers in Psychology*, *7*, 1052. https://doi.org/10.3389/fpsyg.2016.01052

Hendryckx, C., Guillaume, M., Beuel, A., Van Rinsveld, A., & Content, A. (2021). Mutual influences between numerical and non-numerical quantities in comparison tasks. *Quarterly Journal of Experimental Psychology*, *74*(5), 843–852. https://doi.org/10.1177/1747021820981876

Hyde, D. C. (2011). Two Systems of Non-Symbolic Numerical Cognition. *Frontiers in Human Neuroscience*, *0*. https://doi.org/10.3389/fnhum.2011.00150

Inglis, M., & Gilmore, C. (2014). Indexing the approximate number system. *Acta Psychologica*, *145*, 147–155. https://doi.org/10.1016/j.actpsy.2013.11.009

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*, *106*(25), 10382–10385. https://doi.org/10.1073/pnas.0812142106

Kadosh, R. C., & Walsh, V. (2009). Numerical representation in the parietal lobes: Abstract or not abstract? *Behavioral and Brain Sciences*, *32*(3–4), 313–328. https://doi.org/10.1017/S0140525X09990938

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The Discrimination of Visual Number. *The American Journal of Psychology*, *62*(4), 498. https://doi.org/10.2307/1418556

Leibovich, T., Katzin, N., Harel, M., & Henik, A. (2017). From "sense of number" to "sense of magnitude": The role of continuous magnitudes in numerical cognition. *Behavioral and Brain Sciences*, *40*. https://doi.org/10.1017/S0140525X16000960

Leibovich, T., Katzin, N., Salti, M., & Henik, A. (2017). Toward an integrative approach to numerical cognition. *Behavioral and Brain Sciences*, *40*. https://doi.org/10.1017/S0140525X17000619

Lemaire, P., & Lecacheur, M. (2007). Aging and Numerosity Estimation. *The Journals of Gerontology: Series B*, *62*(6), P305–P312. https://doi.org/10.1093/geronb/62.6.P305

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

Lindskog, M., Winman, A., Juslin, P., & Poom, L. (2013). Measuring acuity of the approximate number system reliably and validly: The evaluation of an adaptive test procedure. *Frontiers in Psychology*, *4*, 510. https://doi.org/10.3389/fpsyg.2013.00510

Lourenco, S. F., & Aulet, L. S. (2022). *A Theory of Perceptual Number Encoding*. PsyArXiv. https://doi.org/10.31234/osf.io/p7rqn

Lourenco, S. F., & Bonny, J. W. (2017). Representations of numerical and non-numerical magnitude both contribute to mathematical competence in children. *Developmental Science*, *20*(4), e12418. https://doi.org/10.1111/desc.12418

Marinova, M., Fedele, M., & Reynvoet, B. (2021). Weighted numbers. *The Behavioral and Brain Sciences*, *44*, Article e196. https://doi.org/10.1017/S0140525X21001059

McCrink, K., Dehaene, S., & Dehaene-Lambertz, G. (2007). Moving along the number line: Operational momentum in nonsymbolic arithmetic. *Perception & Psychophysics*, *69*(8), 1324–1333. https://doi.org/10.3758/BF03192949

McCrink, K., & Wynn, K. (2004). Large-Number Addition and Subtraction by 9-Month-Old Infants. *Psychological Science*, *15*(11), 776–781. https://doi.org/10.1111/j.0956-7976.2004.00755.x

Mou, Y., & vanMarle, K. (2014). Two core systems of numerical representation in infants. *Developmental Review*, *34*(1), 1–25. https://doi.org/10.1016/j.dr.2013.11.001

Nieder, A. (2016). The neuronal code for number. *Nature Reviews Neuroscience*, *17*(6), Article 6. https://doi.org/10.1038/nrn.2016.40

Odic, D., & Starr, A. (2018). An Introduction to the Approximate Number System. *Child Development Perspectives*, *12*(4), 223–229. https://doi.org/10.1111/cdep.12288

Pekár, J., & Kinder, A. (2020). The interplay between non-symbolic number and its continuous visual properties revisited: Effects of mixing trials of different types. *Quarterly Journal of Experimental Psychology*, *73*(5), 698–710. https://doi.org/10.1177/1747021819891068

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences*, *14*(12), 542–551. https://doi.org/10.1016/j.tics.2010.09.008

Piazza, M., Pica, P., Izard, V., Spelke, E. S., & Dehaene, S. (2013). Education Enhances the Acuity of the Nonverbal Approximate Number System. *Psychological Science*, *24*(6), 1037–1043. https://doi.org/10.1177/0956797612464057

Pickering, J., Adelman, J. S., & Inglis, M. (2023). *Are Approximate Number System representations numerical?* https://doi.org/10.5964/jnc.8553']

Pomè, A., Thompson, D., Burr, D. C., & Halberda, J. (2021). Location- and object-based attention enhance number estimation. *Attention, Perception, & Psychophysics*, *83*(1), 7–17. https://doi.org/10.3758/s13414-020-02178-w

Price, G. R., Palmer, D., Battista, C., & Ansari, D. (2012). Nonsymbolic numerical magnitude comparison: Reliability and validity of different task variants and outcome measures, and their relationship to arithmetic achievement in adults. *Acta Psychologica*, *140*(1), 50–57. https://doi.org/10.1016/j.actpsy.2012.02.008

Qu, C., DeWind, N. K., & Brannon, E. M. (2022). Increasing entropy reduces perceived numerosity throughout the lifespan. *Cognition*, *225*, 105096. https://doi.org/10.1016/j.cognition.2022.105096

Qu, C., Szkudlarek, E., & Brannon, E. M. (2021). Approximate multiplication in young children prior to multiplication instruction. *Journal of Experimental Child Psychology*, *207*, 105116. https://doi.org/10.1016/j.jecp.2021.105116

Roquet, A., & Lemaire, P. (2019). Strategy variability in numerosity comparison task: A study in young and older adults. *Open Psychology*, *1*(1), 152–167. https://doi.org/10.1515/psych-2018-0011

Rugani, R., Loconsole, M., Simion, F., & Regolin, L. (2020). Individually distinctive features facilitate numerical discrimination of sets of objects in domestic chicks. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-73431-3

Rumbaugh, D. M., Savage-Rumbaugh, S., & Hegel, M. T. (1987). Summation in the chimpanzee (Pan troglodytes). *Journal of Experimental Psychology: Animal Behavior Processes*, *13*, 107–115. https://doi.org/10.1037/0097-7403.13.2.107

Sasanguie, D., Defever, E., Maertens, B., & Reynvoet, B. (2014). The Approximate Number System is not Predictive for Symbolic Number Processing in Kindergarteners. *Quarterly Journal of Experimental Psychology*, *67*(2), 271–280. https://doi.org/10.1080/17470218.2013.803581

Sasanguie, D., De Smedt, B., & Reynvoet, B. (2017). Evidence for distinct magnitude systems for symbolic and non-symbolic number. *Psychological Research*, *81*(1), 231–242. https://doi.org/10.1007/s00426-015-0734-1

Schneider, M., Beeres, K., Coban, L., Merz, S., Schmidt, S. S., Stricker, J., & Smedt, B. D. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*, *20*(3), e12372. https://doi.org/10.1111/desc.12372

Sella, F., Lanfranchi, S., & Zorzi, M. (2013). Enumeration skills in Down syndrome. *Research in Developmental Disabilities*, *34*(11), 3798–3806. https://doi.org/10.1016/j.ridd.2013.07.038

Shutts, K., Condry, K. F., Santos, L. R., & Spelke, E. S. (2009). Core knowledge and its limits: The domain of food. *Cognition*, *112*(1), 120–140. https://doi.org/10.1016/j.cognition.2009.03.005

Sixtus, E., Krause, F., Lindemann, O., & Fischer, M. H. (2023). A sensorimotor perspective on numerical cognition. *Trends in Cognitive Sciences*, *27*(4), 367–378. https://doi.org/10.1016/j.tics.2023.01.002

Smets, K., Sasanguie, D., Szücs, 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. https://doi.org/10.1080/20445911.2014.996568

Spelke, E. S. (2022). *What Babies Know: Core Knowledge and Composition Volume 1*. Oxford University Press.

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89–96. https://doi.org/10.1111/j.1467-7687.2007.00569.x

Strickland, B. (2017). Language Reflects “Core” Cognition: A New Theory About the Origin of Cross-Linguistic Regularities. *Cognitive Science*, *41*(1), 70–101. https://doi.org/10.1111/cogs.12332

Stricker, L. K. (2021). *Mental Imagery und geschätzte Nummerositäten – Die Rolle von räumlichen Anordnungshinweisen bei einer cross-modalen Schätzaufgabe*. Unpublished Bachelorthesis, Justus-Liebig-University Giessen.

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 Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00444

Testolin, A., Dolfi, S., Rochus, M., & Zorzi, M. (2020). Visual sense of number vs. Sense of magnitude in humans and machines. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-66838-5

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

Tokita, M., & Ishiguchi, A. (2016). Precision and Bias in Approximate Numerical Judgment in Auditory, Tactile, and Cross-modal Presentation. *Perception*, *45*(1–2), 56–70. https://doi.org/10.1177/0301006615596888

Tomlinson, R. C., DeWind, N. K., & Brannon, E. M. (2020). Number sense biases children's area judgments. *Cognition*, *204*, 104352. https://doi.org/10.1016/j.cognition.2020.104352

Vallortigara, G. (2012). Core knowledge of object, number, and geometry: A comparative and neural approach. *Cognitive Neuropsychology*, *29*(1–2), 213–236. https://doi.org/10.1080/02643294.2012.654772

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483–488. https://doi.org/10.1016/j.tics.2003.09.002

Wang, J. (Jenny), Halberda, J., & Feigenson, L. (2017). Approximate number sense correlates with math performance in gifted adolescents. *Acta Psychologica*, *176*, 78–84. https://doi.org/10.1016/j.actpsy.2017.03.014

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal Counting in Humans: The Psychophysics of Number Representation. *Psychological Science*, *10*(2), 130–137. https://doi.org/10.1111/1467-9280.00120

Ziegler, M. C., & Drewing, K. (2022). Get in touch with numbers – an approximate number comparison task in the haptic modality. *Attention, Perception, & Psychophysics*. https://doi.org/10.3758/s13414-021-02427-6

Ziegler, M. C., Stricker, L. K., & Drewing, K. (2023). The role of spatial information in an approximate cross-modal number matching task. *Attention, Perception, & Psychophysics*. https://doi.org/10.3758/s13414-023-02658-9

Zorzi, M., & Testolin, A. (2018). An emergentist perspective on the origin of number sense. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1740), 20170043. https://doi.org/10.1098/rstb.2017.0043

II Publications

Attention, Perception, & Psychophysics (2022) 84:943–959
https://doi.org/10.3758/s13414-021-02427-6

Get in touch with numbers – an approximate number comparison task in the haptic modality

Marco Carlo Ziegler[1] · Knut Drewing[2]

Accepted: 8 December 2021

Abstract

The Approximate Number System (ANS) is conceptualized as an innate cognitive system that allows humans to perceive numbers of objects or events (>4) in a fuzzy, imprecise manner. The representation of numbers is assumed to be abstract and not bound to a particular sense. In the present study, we test the assumption of a shared cross-sensory system. We investigated approximate number processing in the haptic modality and compared performance to that of the visual modality. We used a dot comparison task (DCT), in which participants compare two dot arrays and decide which one contains more dots. In the haptic DCT, 67 participants had to compare two simultaneously presented dot arrays with the palms of their hands; in the visual DCT, participants inspected and compared dot arrays on a screen. Tested ratios ranged from 2.0 (larger/smaller number) to 1.1. As expected, in both the haptic and the visual DCT responses similarly depended on the ratio of the numbers of dots in the two arrays. However, on an individual level, we found evidence against medium or stronger positive correlations between "ANS acuity" in the visual and haptic DCTs. A regression model furthermore revealed that besides number, spacing-related features of dot patterns (e.g., the pattern's convex hull) contribute to the percept of numerosity in both modalities. Our results contradict the strong theory of the ANS solely processing number and being independent of a modality. According to our regression and response prediction model, our results rather point towards a modality-specific integration of number and number-related features.

Keywords Multisensory processing · Haptics · Visual perception

Introduction

To be able to perceive and process numbers of events or objects, humans as well as other animals require a neuronal grounding. A commonly accepted idea is that humans are equipped with an innate number sense that allows processing of numerosity (Butterworth, 2010; Dehaene, 2011; Feigenson et al., 2004; Piazza, 2010). Numerosity (also known as cardinality) is defined by Nieder (2016) as the number of countable items in a given set (in contrast to, e.g., discrete numerical symbols, such as Hindu-Arabic numerals). There is an ongoing debate of the exact definition of the number sense and what abilities it includes (Jordan et al., 2006; Szűcs & Myers, 2017). One established theory, the two-component model, distinguishes the number sense into one system that processes small nonsymbolic numbers of distinct individuals that fall into subitizing range (i.e., smaller than five items) and another system for larger numbers beyond four (Dehaene, 2011; Feigenson et al., 2004; Hyde, 2011; Mou & vanMarle, 2014; Olsson et al., 2016; De Smedt et al., 2013). The latter one is often denoted as the Approximate Number System (ANS), a cognitive system that enables people "to represent quantities as imprecise, noisy mental magnitudes without verbal counting or numerical symbols" (Park & Brannon, 2013, p. 1). A signature of the ANS is ratio dependence (e.g., when humans compare two nonsymbolic numerosities; Brannon & Merritt, 2011). Performance in such comparison tasks is determined by the ratio of the two numbers and not their numerical difference. For instance, the comparison performance between 5 and 6 dots is similar to that between 10 and 12 dots, and it is faster and more accurate than that between 7 and 8 dots (Halberda et al., 2008). Ratio dependence results from fuzzy

---

✉ Marco Carlo Ziegler
marco.c.ziegler@psychol.uni-giessen.de

[1] University of Giessen, Department of Psychological Assessment, Giessen, Germany

[2] University of Giessen, Department of Experimental Psychology - HapLab, Giessen, Germany

mental representations of larger numbers and is consistent with Weber´s Law (Brannon & Merritt, 2011; Feigenson et al., 2004; Szkudlarek & Brannon, 2017). Another core assumption of the ANS is that numerosity is represented in an abstract format that is shared across sensory modalities (Barth et al., 2005; Feigenson et al., 2004).

Whereas ratio dependence of approximate numerosity has been demonstrated repeatedly (Feigenson et al., 2004; Mou & vanMarle, 2014), the assumption of a truly abstract representation of number is topic of an ongoing debate (Anobile et al., 2018; Anobile et al., 2016; Barth et al., 2003; Barth et al., 2005; Gebuis et al., 2016). In the present study, we examined the ability to compare approximate numerosities in the visual and the haptic modality and investigated their commonalities and differences.

The ANS theory suggests that of any given stimulus material (e.g., a dot pattern) numerosity gets extracted by filtering out the irrelevant confounding features of the stimuli (like varying diameter of the dots) and then accumulating the numerosity (Brannon & Merritt, 2011; Dehaene, 2011; Dehaene & Changeux, 1993; Gebuis et al., 2016), implying abstract representations that are independent of the used sensory modality. Barth et al. (2003) stated that a truly abstract number sense should be robust against variation of stimulus presentation formats and actual presented stimulus material. Comparing performances of number estimation tasks between different modalities as well as cross-modal paradigms are a straightforward way to examine these assumptions.

Research on larger numerosities has mainly been dedicated to the visual modality whereas other modalities have not (yet) received similar attention. At least a few studies have utilized auditory stimulus material, e.g., tone sequences in audiovisual habituation paradigms with newborns (Izard et al., 2009), or sequential dot and tone sequences in same–different or arithmetic tasks (Barth et al., 2003; Barth et al., 2005). The assumption of a system that is capable of extracting numerosity independently of the sensory input has been addressed by Barth et al. (2003). They distinguished between "not modality specific" representation of number and a modality specific "perceptual representation" that also might include stimulus specific influences. They argued that if a numerosity representation is specific to a modality, cross-modal numerosity comparisons must fail because modality specific percepts will not be compatible. In a series of experiments, Barth et al. (2003) tested participants in conditions varying numerosity task formats (temporal/spatial) along with sensory modality (within-modality/cross-modal): Discrimination of numerosity did not differ significantly between within-modality conditions and cross-modal conditions when both stimuli in a trial were following the same task format (within-modality: two auditory or two visual sequences; cross-modal: comparing an auditory to a visual sequence). Also, discrimination performance within a single modality did not differ when using different task formats (comparing visual sequences to spatial visual dot arrays) as compared with using one task format (two visual sequences or two visual arrays). The authors concluded that there must be a truly amodal representation of number, which is neither specific to a modality nor a task format, because in cross-modal and cross-task conditions participants' performance did not fall behind the performance of their weaker within-modal and within-task condition, respectively.

However, in the same study by Barth et al. (2003), participants performed significantly worse in a combined cross-modal cross-task condition (comparing a visual dot array to tones in a sequence) than in either of the two corresponding within-modal same-task conditions (two visual dot arrays or two auditory sequences). This seems to indicate that the representation of number is at least not completely amodal, and that to some extent perceptual cues (conveyed by task formats) influence the percept of number. Similarly, Anobile et al. (2018) found that individual approximate numerosity performance hardly correlated between a spatial visual and a temporal auditory task (at least in adults), whereas individual performance within the temporal visual and auditory tasks showed substantial correlations in children and adults (Anobile et al., 2018). That is, differences in sensory modality and task format resulted in a lack of correlation in task performance within adult participants, which is strong evidence against an amodal numerosity representation that is completely shared between the two tasks. However, in the study by Anobile et al. (2018), it is not entirely clear how far the lack of correlation leads back to differences in the task format or in the sensory modalities.

Tokita et al. (2013) tried to isolate modality-specific factors influencing numerosity estimation by defining a stricter standardization for numerosity tasks in different sensory modalities. The authors used strictly parallel sequential visual and auditory task formats. Participants compared dots or tone sequences within a modality as well as in a cross-modal condition. Participants' performances (Weber fractions) showed substantial differences between the visual and auditory modality (Tokita et al., 2013). The authors concluded this to be a contradictory result to the assumption of a modality-independent numerical representation system, since a truly abstract system processing numerosity independently would not allow for performance differences. In a subsequent study, Tokita and Ishiguchi (2016) extended this line of research by evaluating participants' performance in comparing tone sequences and sequences of vibrations (via vibro-tactile devices). Here, on the other hand, they found no significant differences in participants' performance between the two modalities, which they considered to be expected as the auditory and the haptic modality are more alike than are the visual and the auditory domain (Tokita & Ishiguchi, 2016).

Overall, results are mixed regarding the question of a modality-independent representation of numerosity. There are also differences in the interpretation of results (Gebuis et al., 2016), which partly relate to different understandings of the ANS concept following either a strong interpretation that the ANS solely extracts number, or a broader interpretation allowing that additionally other factors play some role. A direct comparison of results between studies is further complicated by a wide variety of measurement approaches for approximate numerosity (different instructions, stimulus ranges, presentation formats, etc.).

In the present study, we aim for a straightforward test to investigate the question if number representation is abstract and grounded in a shared modality-independent system, as well as for transparent data interpretation in terms of stronger and broader ANS concepts. The present study uses parallel tasks formats in the haptic and the visual modality (simultaneous presentation of two spatial dot arrays; cf. Tokita et al., 2013) and includes correlations of individual performances as a strong measure of shared representations (cf. Anobile et al., 2018). Surprisingly, the haptic modality has been left almost untouched for investigating these questions. However, spatially separated dot comparison tasks (a.k.a. “paired”-paradigm; Dietrich et al., 2015; Inglis & Gilmore, 2014), can be designed for the haptic modality—as it is the case for the visual but not the auditory modality. Therefore, a dot comparison task (DCT) format in the haptic modality can be implemented fairly similarly to the format that is typically used in the visual modality. From a theoretical perspective, this method, where participants have to decide which one of two spatially separated dot arrays contains more elements (Dietrich et al., 2015), is preferred when assessing a person’s ANS’s acuity, because this paradigm requires the least additional cognitive processes when perceiving numerosity (Dietrich et al., 2015). Pioneer work in approximate numerosity estimation tasks in the haptic modality was done by Ginsburg and Pringle (1988), who, however, used a direct numerosity estimation task and not a paradigm with simultaneously presented separated dot arrays. To our knowledge, the only adaptation of a haptic DCT with simultaneously presented and spatially separated arrays has been reported by Gimbert et al. (2016). They presented two stimulus sets (rectangular pads with raised dots) to 147 children and asked them to explore the pads simultaneously via active touch and to compare dot numerosity. Pairs of pads represented ratios from 1.1 (hardest) to 3.0 (easiest). The children additionally performed a visual DCT. The children were able to compare approximate quantities in the haptic DCT and showed typical and similar ratio dependent effects in the visual as well as in the haptic DCT. Gimbert et al. (2016) concluded that a shared mechanism (approximate number processing) accounted for their results.

However, this conclusion needs additional testing, because it is mainly based on the observation that the average haptic and visual performance both increase with higher ratios. As a number of different mechanisms would predict better performance with higher ratios, the shared average performance in visual and haptic tasks is not sufficient evidence for a common mechanism. A whole line of recent research has identified mechanisms that contribute to participants’ numerosity estimates, such as the convex hull or area density of a dot array (Bertamini et al., 2016; Clayton et al., 2015; Tomlinson et al., 2020). Our approach here is to estimate such factors that might additionally contribute to the number percept, and to find their relative weight in the haptic and visual response process. Thus, we investigate whether number is indeed the strongest influence to account for the results. We use an approach suggested by DeWind et al. (2015). DeWind et al. (2015) describe different orthogonal factors (number, spacing, and size) that contribute to the percept and the participant’s final response in a DCT (details in data analysis). Further, we aim to assess individual visual and haptic performance in DCTs rather than average group performance alone. The individual then serves as an additional factor besides ratio that influences performance and can be used to test for a shared mechanism: If the ANS is cross-modally shared, differences in individual performance should covary between the two sensory modalities. Moreover, we also want to advance some stimuli and design choices in the study of Gimbert et al. (2016), which may have been problematic.

First, we implemented sufficient spacing between dots in the haptic DCT (>11 mm based on Craig & Lyle, 2001) to allow participants to segregate dots as singular entities rather than perceiving them as an essentially not-countable texture. Second, Gimbert et al. (2016) seem to have used the same limited number of dot stimuli for every child in their study, which risks creating confounds and biases related to the specifically used dot configurations. In the present study, we hence varied the dot patterns used between sessions and participants. Third, we collect sufficient data to be able to report every tested ratio separately rather than an aggregation of different ratios into categories (“ratio bins”), and to allow for a basic check for psychometric features (e.g., a reliability estimation) in our novel implementation of a haptic DCT. Finally, we also calculated mean percentage correct and Weber fractions as ANS acuity indicators (omitted in Gimbert et al., 2016), which can be compared between modalities (see Anobile et al., 2018), to investigate if there is a shared underlying system that processes the same feature.

In the present study, we presented a visual and a haptic DCT to 71 individuals. Our haptic and visual DCT share all essential features, which makes them comparable to the extent that if number is the primary source for the percept, ANS acuity in both tasks should be highly correlated. These include the general paired paradigm (simultaneous spatial dot arrays) and identical dot ratios. Obviously, as was the case in the previous studies, also some modality-specific adaptations were necessary, which we report in detail below. We evaluated

a) whether adult participants express typical ratio dependent performance in both tasks,
b) the relative weight of factors that lead to the percept of numerosity by means of a regression model of numerosity perception for each modality,
c) the test–retest reliability of the haptic DCT and the visual DCT, and
d) whether participants' individual and group-wise performances correlate between the haptic and the visual DCT.

Methods

Participants

Using G-Power (Version 3.1; Faul et al., 2009), we estimated a required sample size of 64 participants to achieve a power of .80 (α = .05) for a single-sided medium sized effect (ρ = .3) in the correlation of performance indices between sensory modalities. We recruited 71 participants (49 females, 22 males, mean age 24 years), 67 met the inclusion criteria and remained for the data analysis. All participants were healthy and had normal or corrected to normal vision. None of the participants had any injuries, unusual keratinization or scar tissue in the palm of their hands. Participants got course credit or eight euro per hour as compensation for their efforts. Participants gave written informed consent before enrolling in the study. All procedures and methods were consistent with the World Medical Association Declaration of Helsinki (2013) and were approved by the local ethics committee of Fachbereich 06 of the Justus-Liebig-University Gießen (LEK-FB 06).

Design, setups, and procedure

In the experiment, each participant performed two variants of dot comparison tasks (DCTs), a visual and a haptic DCT. They had to decide which of two dot arrays contained a higher number of dots. We varied the number ratio of the presented dot arrays in the following steps: 2.00, 1.33, 1.20, 1.14, and 1.11 with a number of dots ranging from 5 to 20 per stimulus (cf. Halberda et al., 2008). We conducted each dot comparison in a "single" and a "double" variant ("version"). The double variant contained twice as many dots per stimulus compared with the single variant (Table 1). Furthermore, the more numerous dot patterns were equally frequent on the left and the right side of the screen. Our primary measure for individual ANS acuity was the individual Weber fraction (*w*) per modality condition (DeWind et al., 2015).

Table 1 Complete list of all dot comparisons and resulting number ratios for both visual and haptic DCTs including the additional applied control conditions for the visual DCT

Comparison		Number ratio	Visual condition
10 vs. 5	(single)	2.00	DSC / SAC
5 vs. 10	(single)	2.00	DSC / SAC
20 vs. 10	(double)	2.00	DSC / SAC
10 vs. 20	(double)	2.00	DSC / SAC
6 vs. 8	(single)	1.33	DSC / SAC
8 vs. 6	(single)	1.33	DSC / SAC
12 vs. 16	(double)	1.33	DSC / SAC
16 vs. 12	(double)	1.33	DSC / SAC
6 vs. 5	(single)	1.20	DSC / SAC
5 vs. 6	(single)	1.20	DSC / SAC
12 vs. 10	(double)	1.20	DSC / SAC
10 vs. 12	(double)	1.20	DSC / SAC
8 vs. 7	(single)	1.14	DSC / SAC
7 vs. 8	(single)	1.14	DSC / SAC
16 vs. 14	(double)	1.14	DSC / SAC
14 vs. 16	(double)	1.14	DSC / SAC
10 vs. 9	(single)	1.11	DSC / SAC
9 vs. 10	(single)	1.11	DSC / SAC
20 vs. 18	(double)	1.11	DSC / SAC
18 vs. 20	(double)	1.11	DSC / SAC

Note. DSC = average dot size control; SAC = surface area control.

Data collection consisted of two separate sessions. Each session took place in a quiet laboratory room in the faculty of psychology at the Justus-Liebig-University Gießen. We conducted the haptic and the visual DCT once per session, respectively. Haptic and visual DCTs were performed consecutively within each session. In the second session, we reversed the order of the DCTs for the participant. Whether the haptic or the visual task was first within a session was balanced between participants. We scheduled the sessions one week apart from each other (7 days test–retest interval). The visual DCT took about 30 minutes. The haptic DCT took about 50 minutes. In each DCT, the experimenter monitored the participant over the course of the experiment. The participants could pause the experiments any time to regain focus on the task.

Visual DCT setting

Participants sat at a table with 90 cm distance to a Dell P2213LED Display (22 inches display diagonal) that we

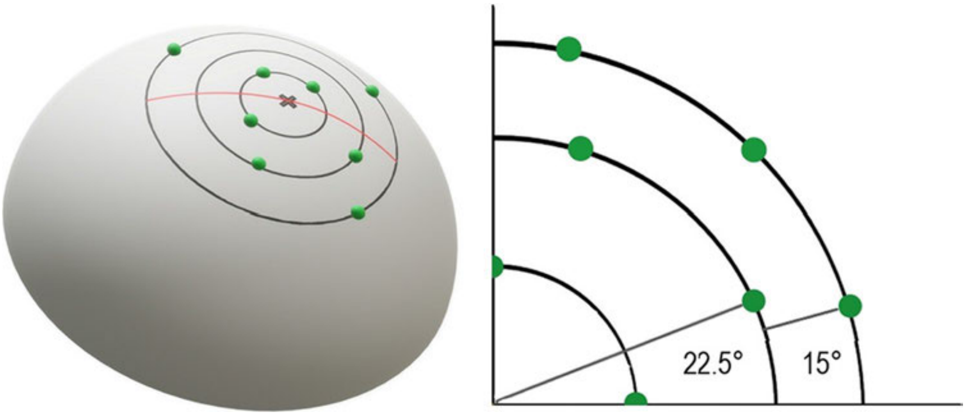


Fig. 1 Left side: Styrofoam hemisphere with exemplary dots (pins) on three radial orbits. Right side: Quadrant with all possible slots for a pin. Position of the slots in degrees from the origin for one quadrant. Slots in the first orbit are perpendicular to the origin

used as an output device. The display ran on a resolution of 1680 × 1050 px (refresh rate 60 Hz). An ordinary QWERTZ keyboard was located in front of the participant as an input device. In each trial, a custom-made C++ program presented two dot patterns on the screen, one on the left side and one on the right side. The dots of the left pattern were placed randomly in a range of 1/8 to 3/8 in width and 1/4 to 3/4 in height of the screen size. The dots of the right pattern were placed randomly in a range of 5/8 to 7/8 in width of the screen size with the same range of height as the left pattern. Dots varied randomly in 40% of diameter around the base diameter size of 10 pixel per dot, which is between ~0.33° and 0.65° diameters in visual angle (participants visual angle: ~17°).

We presented dot patterns in two different ways that allowed us to control for different possible confounding variables (see Halberda et al., 2008). In the average dot size control condition (DSC, 50% of trials), the dots in both patterns had on average the same diameter. The disadvantage of this otherwise favorable feature is that the displayed dot patterns differ in their cumulative surface area on the screen (more pixels for the more numerous set). Therefore, the other 50% of all trials were surface area controlled (SAC), which ensures that the total pixel count of the two dot patterns on the screen were equal.

Visual DCT procedure

Each trial in the visual DCT started with a white fixation cross on black background. When the participant manually initiated the trial by a button press, the fixation cross vanished and two white dot patterns appeared upon black background. After 200 ms, the dot patterns disappeared and the whole screen got flushed with a white mask (300 cd/m^2) to inhibit possible visual aftereffects. Participants responded with the correspondent arrow keys (left or right) which dot pattern contained a higher number of dots. Then the next trial started. Participants completed 40 practice trials in each of the sessions and 720 experimental trials subsequently (18 repetitions of blocks of 40 trials = 5 ratios × 2 versions × 2 visual control × 2 side of higher number). In each repetition, we shuffled the order of trials randomly.

Haptic DCT setting

Participants sat at a table behind an opaque curtain, which prevented the participant from seeing the stimuli during the whole experiment. The experimenter sat opposite to the participant on the other side of the curtain and exchanged the stimuli according to a predefined random trial list. We presented dot patterns in form of pins on Styrofoam hemispheres (industrial map pins of 5 mm in diameter each; see Fig. 1). The hemispheres were 150 mm in diameter and 75 mm in height. We coated them with jersey fabric to ensure a plain and pleasant surface. We marked three radial orbits (radii 15, 30, and 45 mm) onto the fabric in order to systematically organize the pins and capture their position. We centered the orbits on the "pinnacle" of the hemispheres. The inner orbit contained four slots, the intermediate eight slots, and the outer 12 slots for pins. The slot arrangement asserted that pins never formed a straight line, and that the distance between pins was always bigger than 11 mm. We conducted pilot tests ($N = 5$) to find the correct spacing, which allows to accurately individuate every single pin and minimizing the occupied area for convenient exploration. We found that a spacing of 11 mm between dots allowed a participant to reliably individuate a single pin in the palm of the hand (i.e., no errors without time limit). This result is consistent with that of Craig and Lyle (2001) who showed that different grating spaces are reliably discriminated when the spacing is set to 11 mm.

Table 2 Descriptive statistics of the convex hull ratio for each DCT (visual condition 48,240 trials; haptic condition 10,720 trials)

Convex hull ratio	Min	1st Quantile	*Md*	*M*	3rd Quantile	Max	*SD*
Visual DCT	0.01	0.60	0.76	0.72	0.88	1.00	0.17
Haptic DCT	0.18	0.68	0.83	0.78	0.91	1.00	0.20

We preconfigured the stimuli for each participant and each session (i.e., we filled the slots on the hemispheres with pins according to a custom written C++ program that randomly chose positions on the hemisphere; two per 5 ratios × 2 versions = 20 hemispheres). In all single version hemispheres, we placed pins to slots of the inner and the intermediate orbit. In the double version hemispheres, we used all three orbits to enhance stimulus variety. We presented the hemispheres of each of the 10 pairs in two different orientations (0° and 120°) in order to increase the diversity of perceived pin configurations. In the haptic task, all pins were of the same size, so the haptic task can be considered an analogue to the visual dot-size control (DSC) condition.

Haptic DCT procedure

A haptic DCT trial started when the experimenter had placed the two Styrofoam hemispheres for exposure and the participant had stated to be ready. Five seconds of exploration time started when the palms of their hands touched the two stimuli. We chose the 5 seconds exploration time to ensure that participants can only approximate numerosity, as previously demonstrated in Ginsburg and Pringle (1988). We additionally validated that a participant's counting performance for 10 pins in 5 seconds is at chance level in our own pilot data. The participants examined two hemispheres simultaneously with the palms of their hands. The experimenter rigorously observed the haptic exploration process, monitored the time limit with a stopwatch and announced when the time limit elapsed. After the exploration time expired, the experimenter asked the participant to raise hands and stop exploring the stimuli. The participant responded verbally with "left" or "right" which side contained more dots. If the experimenter had noticed that a participant did not touch all pins on the hemisphere or exceeded the time limit, the trial was repeated later in the experiment. In each session, participants initially examined a sample hemisphere with equipped pins, and then practiced the task in two trials. Afterwards, participants completed 80 experimental trials (5 ratios × 2 versions × 2 orientations × 2 side of higher number × 2 repetitions), which were ordered randomly.

Modality specific DCT adaptations

Our haptic and visual DCT share essential features of the general paired task paradigm, number ratios and the DSC. We, however, also made adaptations to account for particularities of modalities. First, the presentation time of stimuli differs between haptic and visual DCT (cf. Gimbert et al., 2016): Pilot investigations had shown that participants require a substantially longer time to extract the ("numerosity") information with the palms of their hands as compared with the visual DCT, and we hence set haptic exploration time to 5 s. In contrast, extending the presentation time for the visual DCT toward the haptic time in the visual condition would result in participants being able to serially count dots. Thus, to tap similar processes in both tasks, we needed to use different presentation times.

Second, we used different visual dot sizes, but we did not use pins of different size in the haptic condition, nor did we implement a haptic analogue to the visual surface-area control condition. This is because the relation between pin size and haptic stimulation is different than that between dot size and visual stimulation: Haptic stimulation intensity relates to the extent of local skin deformation (Hayward, 2011), meaning that smaller pins will cause more intense stimulation than larger ones, and intensity further depends on pin shape and skin site. Thus, a straightforward transfer of the SAC was not possible, and the effects of different pin sizes would not have been well controlled. Still, the haptic task can be considered analogue to the visual DSC condition.

Third, due to different requirements of the DCTs (e.g., different interdot distances; visual discrimination of dots only requires an interdot distance as low as 0.28 mm), the overall necessary area, which a dot pattern occupies, as well as the overall possible density of the dot patterns, produces different variances in the convex hull ratio between the modalities. The relative convex hull of two dot patterns within a trial could vary from a minimum ratio of 1 to a maximum factor of .012 in the visual DCT and from a minimum ratio of 1 to a maximum factor of .183 in the haptic DCT. Our protocol generated "truly" random dot patterns for both modalities within their area boundaries and protocol restrictions (like interdot distance) resulting in different yet comparable and, due to a large number of trials, balanced range across the convex hull ratio scale (see Table 2).

Last but not least, the number of trials was much higher in the visual as compared with the haptic task, because haptic trials take longer than visual trials. Reducing the number of visual trials to that of haptic trials might have increased parallelism between the tasks, but in our view, a better visual measurement was preferred.

Overall, despite and partly by means of these modality-specific adaptations, our design is well-suited to answer the question of whether the main source of a participant's percept is indeed a number. In addition, our approach utilizes statistical methods proposed by DeWind et al. (2015), which allow us to determine a statistically controlled performance index that mainly reflects the ANS acuity (see Data Analysis section).

Data analysis

For each participant, sensory modality, version, and session, we calculated the mean percentage correct per ratio. We used the mean percentage of correct answers to investigate the ratio dependency and to scrutinize the data for deviance in control conditions and then calculated Weber fractions. We excluded participants' data if their overall mean percentage correct deviated more than three standard deviations from the sample mean in any of the sessions in any modality. Four of the 71 participants met this exclusion criterion. We then aggregated participants' mean percentage correct over sessions, and submitted arcsine transformed (cf. Cohen et al., 2015, pp. 240–241) values into a repeated-measures analysis of variance (ANOVA), with the within-subjects variables ratio, version, and modality. We applied Greenhouse–Geisser correction (Geisser & Greenhouse, 1958) in case of sphericity violations.

To investigate the association between sensory modalities and task reliability, we used Weber fractions (w) as ANS acuity measure (see Equation 3). We calculated group ws for each sensory modality (visual/haptic) collapsing trials over all sessions, separate group ws for trials of the visual control condition DSC and SAC, and individual ws for each participant using trials of the two sessions (in the respective modality) separately to estimate task reliability, and for both sessions combined in order to study cross-modality correlations. Here, we additionally excluded individuals if their w deviated more than three standard deviations from the sample mean of the respective condition. This led to a partial exclusion of participants (up to three) in some visual conditions.

Further, we used the method proposed by DeWind et al. (2015) to quantify the contribution of different stimulus features in a DCT to the responses of participants. The method models the probability of choosing the stimulus that is placed on one particular side (here, the right side), as a function of influence factors that contribute to the percept of number. We pooled the response data over all participants for each modality (separate for the visual controls DSC and SAC). For the response model, influence factors on the number percept are estimated via a regression model in form of regression weights. We adapted Equation 7 in DeWind et al. (2015). The probability to choose the right side is then given by the equation:

$$p(choose\ right) = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{\log_2\left(r_{num}\right) - (\mu)}{\sqrt{2}\beta_{num}{}^{-1}}\right)\right), \quad (1)$$

where μ is given by:

$$\mu = \frac{-\beta_{side} - \beta_{spacing}\log_2\left(r_{spacing}\right)}{\beta_{num}}. \quad (2)$$

In Equations 1 and 2, erf is the Gaussian error function, β_{num} and $\beta_{spacing}$ are the regression weights for the number ratio and spacing ratio, respectively, and β_{side} reflects a bias towards a side. The number ratio r_{num} is calculated by dividing the number of dots on the right side by the number of dots on the left side (log-values then range from -1, which means that on the left side dots are twice as numerous, to 1, which means dots are twice as numerous on the right side). The spacing ratio predictor $r_{spacing}$ is a ratio of sparsity (convex hull divided by number of dots on the respective side) times the convex hull area for the respective sides. Regression parameters β_{num}, $\beta_{spacing}$, β_{side} in Equations 1 and 2 are obtained by fitting a generalized linear model (iteratively reweighted least squares) via a probit link function of the response variable (0 = left side, 1 = right side) with $\log_2$ transformed predictors of number ratio and spacing ratio.

The regression weights themselves indicate a relative contribution of the factor to the number percept and therefore should be identical for each modality, if the exact same processes determined the number percept. To test relevant regression coefficients between the models for significant differences we utilized a z-test procedure proposed by Clogg et al. (1995).

We applied the methods used to estimate the overall model's regression weights also to each individual data set in order to estimate the individual regression coefficient $\beta_{num,\ i}$. From this coefficient, we determined each participant's Weber fraction w_i as a measure of ANS acuity using the following equation:

$$w = \frac{1}{\sqrt{2}\ \beta_{num}}. \quad (3)$$

For a detailed description of how to prepare the predictors and an in-depth explanation of the model, we refer to DeWind et al. (2015). As a brief summary, we use the regression model to quantify the influence of spacing and

Table 3 Relative frequencies of correct trials (*M*) and standard deviations (*SD*) of 67 participants sorted by number ratio, session (t), and visual and haptic DCT

		Ratio									
		2.00		1.33		1.20		1.14		1.11	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
t1											
	Visual	.977	.032	.831	.056	.741	.053	.678	.050	.642	.046
	Haptic	.991	.027	.876	.095	.756	.116	.680	.139	.665	.101
t2											
	Visual	.979	.042	.846	.067	.742	.056	.684	.058	.647	.050
	Haptic	.997	.013	.880	.096	.736	.129	.704	.134	.666	.129

number to the numerosity percept in both modalities and use resulting individual *w*'s for a correlation of performance across modalities; for reliability estimates and group *w*s for an overall evaluation of performance between modalities.

We used R (Version 3.6.1; R Core Team, 2019) for all data analyses and utilized the ggplot2 library (Wickham, 2016) for data visualization.

Results

Ratio dependency

Table 3 lists the aggregated mean percent correct of all 67 participants sorted by number ratio and session. Figure 2 shows the average percentage correct responses for each modality and number ratio, aggregated over both sessions.

The repeated-measures ANOVA (rmANOVA) with the variables number ratio (2.00, 1.30, 1.20, 1.14, and 1.11) and sensory modality (haptic, visual) revealed a significant main effect of ratio, $F(4, 264) = 1367.56, p < .001, \eta_p^2 = .954$. All repeated contrasts of adjacent levels of the variable ratio were significant, all *p*s < .001 (2.00 vs. 1.33, 1.33 vs. 1.20, 1.20 vs. 1.14, 1.14 vs. 1.11) after Bonferroni adjustment. Participants showed in general more correct responses for higher ratios (see Fig. 2). Also, the main effect of modality was significant, $F(1, 66) = 17.81, p < .001, \eta_p^2 = .212$. Participants performed overall better in the haptic DCT compared with the visual DCT, $M_{hap} = 79.5\%$ ($SD = 4.10$) to $M_{vis} = 77.7\%$ ($SD = 3.71$). The Ratio × Modality interaction was significant, $F(4, 264) = 7.39, p < .001, \eta_p^2 = .101$. Interaction contrasts between pairs of successive levels of factor ratio show that the interaction is driven by comparison 1.33 versus 1.20, $F(1, 66) = 9.32, p = .003$, other interaction contrasts turned out not to be significant.

We also compared the "single" and the "double" version for each modality. The rmANOVAs with the variables number ratio (2.00, 1.33, 1.20, 1.14, and 1.11) and version (single, double) revealed a significant main effect of ratio, $F(4, 264) = 404.77, p < .001, \eta_p^2 = .860$ (haptic) and $F(4, 264) = 1751.29, p < .001, \eta_p^2 = .964$ (visual), but not of version neither in the visual, $F(1, 66) = 0.24, p = .629, \eta_p^2 = .004$ ($M_{\text{single}} = 77.6\%, M_{\text{double}} = 77.8\%$), nor in the haptic modality, $F(1, 66) = 1.79, p = .186, \eta_p^2 = .026$ ($M_{\text{single}} = 80.1\%, M_{\text{double}} = 78.9\%$). The Ratio × Version interaction was not significant neither in the visual, $F(4, 264) = 0.96, p = .424, \eta_p^2 = .014$, nor in the haptic modality, $F(4, 264) = 1.76, p = .153, \eta_p^2 = .026$. This indicates ratio dependency because the total amount of dots is not a significant factor.

Furthermore, we conducted three Bonferroni adjusted (corrected for three comparisons) rmANOVAs with one variable always being number ratio and the other variable being the control condition in the visual modality (DSC/SAC), session in the haptic modality or session in the visual modality. Any of these three analyses showed a significant factor number ratio exclusively, $p < .001$, but no significant differences in the respective control condition or session variable, nor interaction effects, except of a significant main effect in the visual control condition DSC versus SAC with $F(1, 66) = 40.35, p < .001, \eta_p^2 = .379$ ($M_{DSC} = 78.7\%$, $SD = 3.57$; $M_{SAC} = 76.6\%$, $SD = 4.29$). Even though differences in the visual controls turned out to be significant, we aggregated the trials of the respective conditions for the correlational analysis because individual ANS acuity indicators (*w*) for SAC and DSC correlated strongly with each other, $r(66) = .871, p < .001$. For the response modeling however, we treated the visual conditions separately to contrast them against the haptic responses and investigate their differences.

Correlational analysis and reliability

The histogram in Fig. 3 shows the distribution of the individual Weber fractions (*w*) per sensory modality fitted over both sessions.

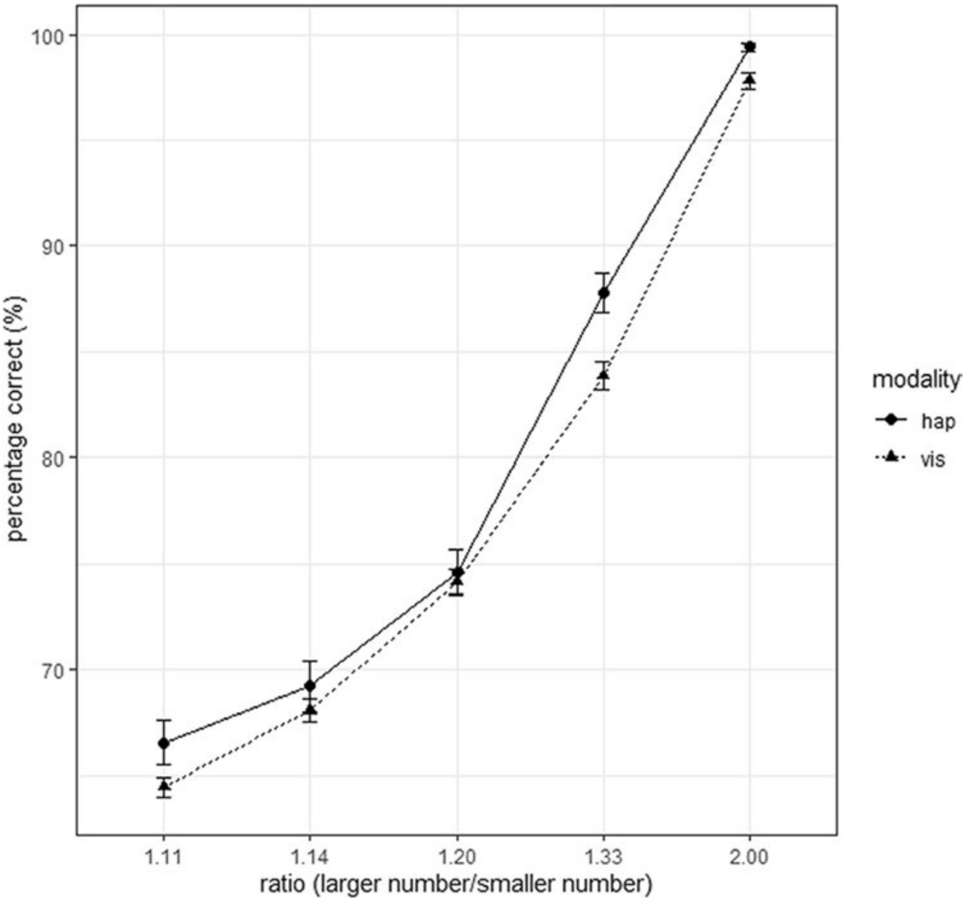


Fig. 2 Ratio-to-correct responses plot for visual and haptic DCTs aggregated over sessions (semi log plot). Error bars represent the standard error of the mean

Table 4 shows the correlations for individual *w*'s between sessions per sensory modality as well as the correlations between the two modalities (trials aggregated over sessions). From intramodal correlations between sessions we calculate the Spearman–Brown adjusted reliability (Brown, 1910) coefficients, because these take into account that the complete haptic and visual tasks (from which we calculate cross-modal correlations) are based on trials from both sessions. Other statistics in Table 4 belong to the nonadjusted values.

Figure 4 shows the scatterplots for each of the three correlations shown in Table 4.

The visual DCT shows with an intersession-correlation of $r(62) = .736$, $p < .001$ an acceptable adjusted task reliability of .848. The correlation coefficient within the haptic DCT shows a smaller but significant association, $r(65) = .251$, $p = .041$, implying a task reliability of .401. The cross-modal performance coefficient does not indicate statistically significant correspondence, $r(63) = -.133$, $p = .290$. It is noteworthy that the confidence interval of the cross-modal correlation does not cover the case of a medium-sized positive correlation effect of $\rho = .3$, rejecting the hypothesis that there is a medium-sized or larger correlation between haptic and visual individual performance. One may argue that the suboptimal haptic reliability may have obscured the true association between modalities. However, when applying a correction for attenuation (Spearman, 1904) in order to estimate correlations

$$\frac{r_{v,h}}{\sqrt{Rel_v} * \sqrt{Rel_h}} \tag{4}$$

without measurement error, the values still reject the hypothesis of a medium-sized or larger true positive correlation (95% CI after applying attenuation correction to its upper and lower borders: −0.449 to 0.018).

Modality specific models of numerosity comparison

We modeled the response behavior of participants (data collapsed over participants) for the visual and the haptic dataset according to Eqs. 1 and 2. This allowed us to calculate the probability to choose the right side for both modalities

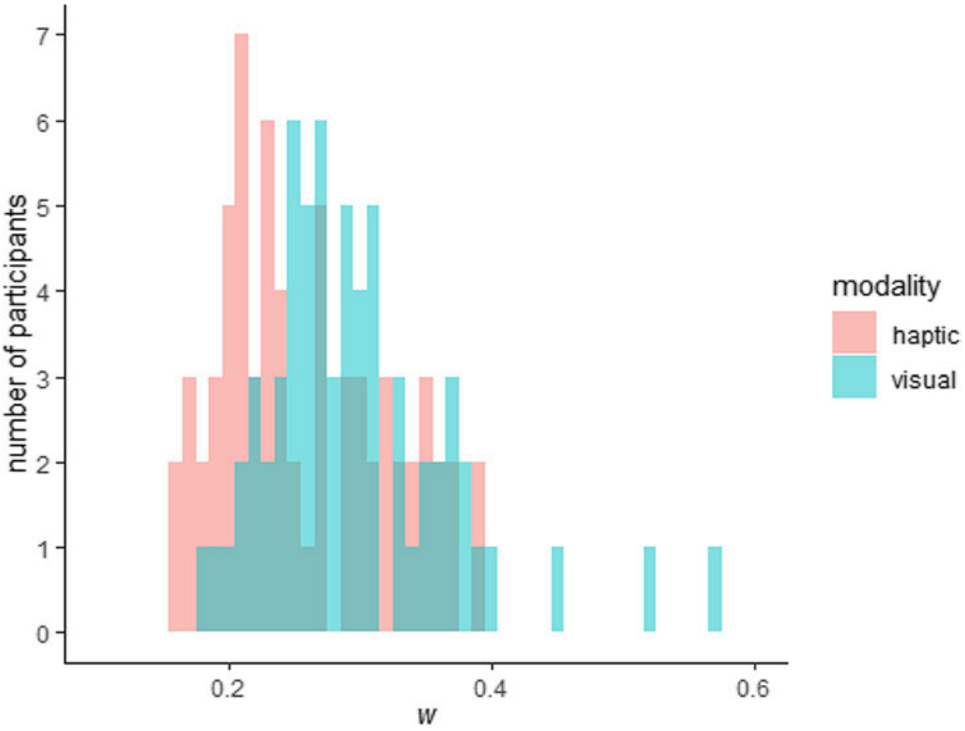


Fig. 3 Distribution of individual *w*s (ANS acuity) in the visual and haptic DCTs. Outlier values are not depicted

Table 4 Correlations of ANS acuity (*w*) across individuals, test statistics, and adjusted task reliability (ATR)

Modality	*r*	*N*	*t*	*df*	CI 95% [lower]	CI 95% [upper]	*P*	*ATR*
haptic–haptic	.251	67	2.088	65	0.011	0.463	.041*	.401
visual–visual	.736	64	8.549	62	0.598	0.831	<.001**	.848
haptic–visual	−.133	65	−1.067	63	−0.365	0.114	.290	–

Note. Intramodal correlations are calculated between first and second session; cross-modal correlations are calculated between modalities (trials aggregated over both sessions, respectively). Spearman–Brown corrected reliability (for the number of trials in the cross-modal correlation) is reported. Statistics calculated for 80 trials.

as a function of number, spacing-related features and side bias. We investigated DSC and SAC controlled trials from the visual DCT separately. Table 5 shows statistics of $\log_2$-transformed predictors.

The results of the visual models (each $n_{\text{visual}} = 48{,}240$ trials) and haptic model ($n_{\text{haptic}} = 10{,}720$ trials) to predict (on group level) the response probabilities to choose the right side dependent on numerosity are depicted in Fig. 5. The model statistics are given in Table 6.The corresponding estimated regression weights are given in Table 7. The statistics show similar and reasonably good pseudo R^2 indices (cf. McFadden, 1977) across all three models. The AICs differ between models, which can be attributed to differences in the total amount of sample trials between modalities.

We compared the regression weights β_{num} and $\beta_{spacing}$ of number ratio (r_{num}) and spacing ratio ($r_{spacing}$), respectively, between the three models (DSC, SAC, haptic) in multiple Bonferroni-adjusted z tests (Clogg et al., 1995) There are no significant differences between the coefficients $\beta_{spacing}$: $z_{haptic/DSC} = -0.116$, $p = .452$; $z_{haptic/SAC} = -0.232$, $p = .492$; and $z_{DSC/SAC} = -0.341$, $p = .488$. However, the coefficient β_{num} was significantly higher for the haptic as compared with the visual models, $z_{haptic/DSC} = 3.244$, $p < .001$, and $z_{haptic/SAC} = 8.716$, $p < .001$, and it was higher for visual DSC as compared with the visual SAC model, $z_{DSC/SAC} = 10.011$, $p < .001$. Correspondingly (cf. Equation 3), the group Weber fraction w (i.e., the estimate for the average ANS acuity) indicates the best ANS acuity for the haptic modality, $w_{haptic} = .28$, and the worst for the visual SAC condition, $w_{DSC} = .30$, $w_{SAC} = .35$.

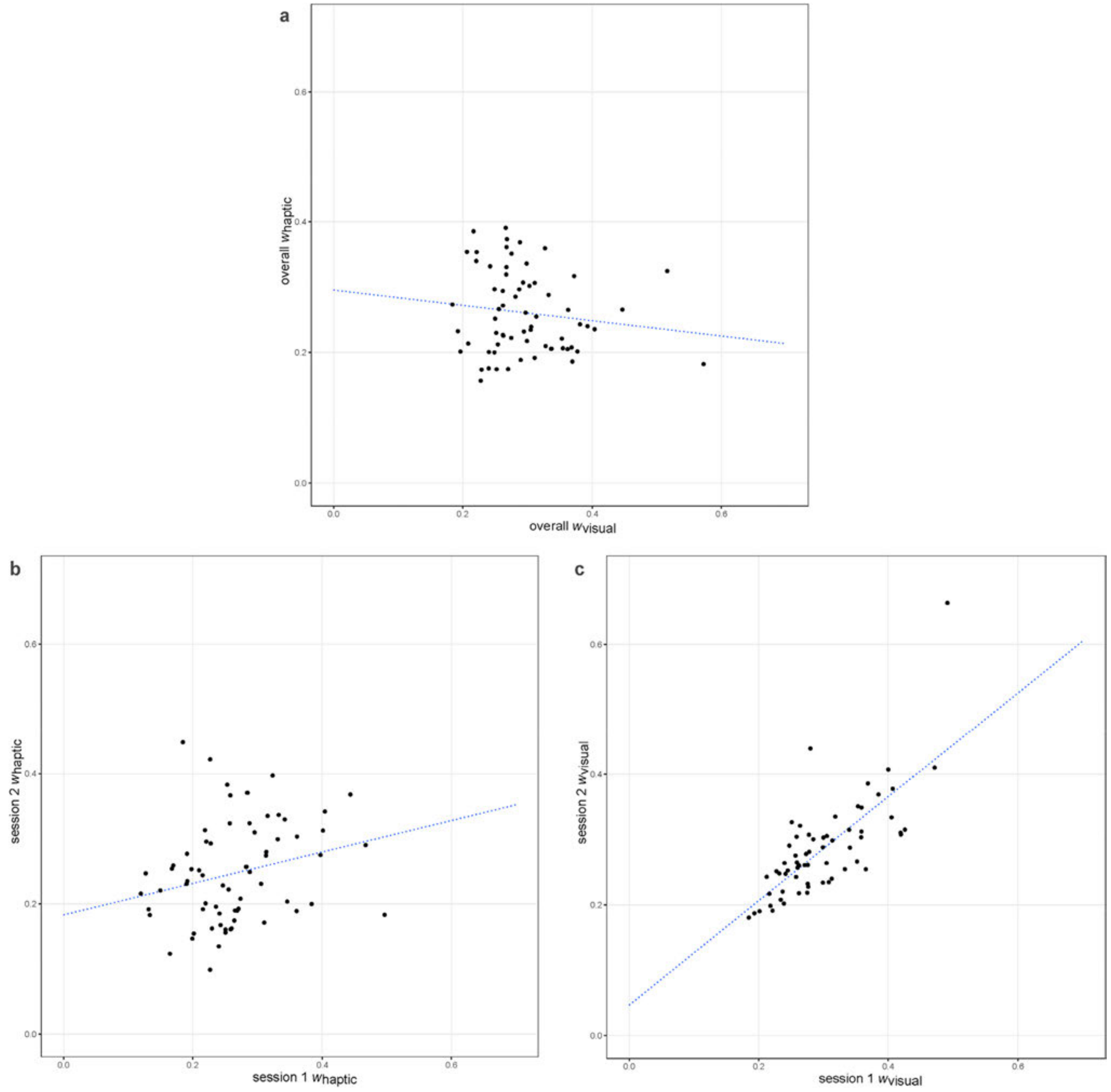


Fig. 4 Scatterplots of individual *w*'s (ANS acuity) for the visual and the haptic DCTs. **a** visual–haptic, **b** haptic t1–t2, and **c** visual t1–t2. Intramodal correlations are calculated between first and second session; cross-modal correlations are calculated between modalities (aggregated over both sessions; t1 = Session 1, t2 = Session 2, respectively)

Discussion

This study investigated nonsymbolic number processing in the haptic and visual modality by means of a dot comparison task. We evaluated if there is evidence for a shared cognitive system (ANS), which processes number independent of the sensory modality. We modeled the performance in the DCT in both modalities on an individual and a group level and evaluated if the measurement is (a) reliable and (b) if performance in both modalities is associated with one another. Our results show that regardless of whether stimuli were examined by touch or by vision, the responses depended on the number ratio of dots between the presented dot patterns, but not on the absolute number of dots. This is a clear indicator of ratio dependency in both modalities. Our group regression model revealed that in the haptic as well as in the visual modality, the ratio of the compared numbers had

Table 5 Descriptive statistics of the $\log_2$-transformed predictors we used for the logistic regression for each modality (visual condition 48,240 trials, haptic condition 10,720 trials)

Predictor (*log_2*)		Min	1st Quantile	*Md*	*M*	3rd Quantile	Max
Haptic							
	r_{num}	−1.00	−0.26	0.00	0.00	0.26	1.00
	$r_{spacing}$	−3.90	−0.39	0.00	0.00	0.39	3.90
Visual$_{DSC}$							
	r_{num}	−1.00	−0.26	0.00	0.00	0.26	1.00
	$r_{spacing}$	−11.74	−0.68	0.00	0.00	0.67	9.94
Visual$_{SAC}$							
	r_{num}	−1.00	−0.26	0.00	0.00	0.26	1.00
	$r_{spacing}$	−9.67	−0.68	0.00	0.00	0.67	9.33

Note. r_{num} refers to number ratio, $r_{spacing}$ to spacing ratio. DSC = average dot sized controlled visual condition; SAC = surface area controlled visual condition.

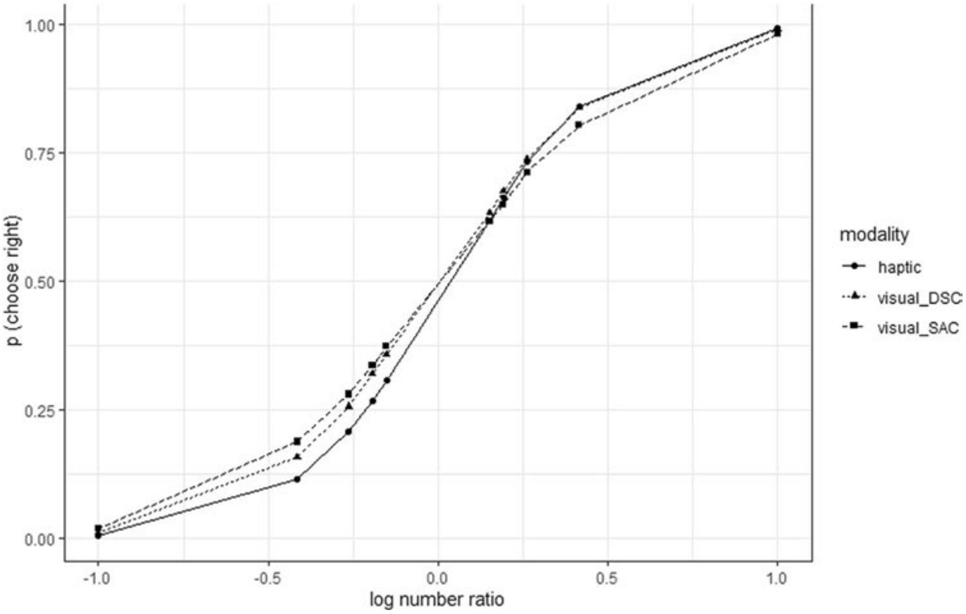


Fig. 5 Probabilities of choosing the right side differentiated by the $\log_2$ ratio of the two compared numerosities (number ratio). Log number ratio of zero indicates that the two dot patterns are equal in number

the primary, major influence on the participants' responses. However, the influence of this predictor slightly but significantly differed between the sensory modalities. Additionally, we found for both modalities a similarly sized smaller influence for less abstract, spacing-related factors (driven by convex hull), which is consistent with previous findings in the visual domain (Clayton et al., 2015). Only in the haptic modality, an additional tendency to choose a side (the left) regardless of the presented stimuli turned out to be a significant predictor but with a rather small effect. These results convincingly show that number ratio is the major influencing factor in both modalities when comparing numerosities of dot arrays in a DCT. However, they do not

Table 6 Model statistics of the logistic regression for each modality (visual condition 48,240 trials, haptic condition 10,720 trials)

Model	AIC	*Null deviance*	pseudo *R*-squared
Haptic	9411.6	14832.8	0.365
Visual$_{DSC}$	42914	66872	0.358
Visual$_{SAC}$	45531	66869	0.319

Note. AIC = Akaike´s Information Criterion. McFadden Pseudo *R*-squared. DSC = average dot sized controlled visual condition; SAC = surface area controlled visual condition.

Table 7 Estimated model coefficients and statistics for coefficient tests against 0 for each modality (visual condition: 48,240 trials; haptic condition: 10,720 trials)

Coefficient		β	Z	SE	p
Haptic					
	β_{num}	2.534	48.39	0.052	<.001***
	$\beta_{spacing}$	0.264	10.38	0.025	<.001***
	β_{side}	−0.101	−6.82	0.015	<.001***
Visual DSC					
	β_{num}	2.349	103.462	0.023	<.001***
	$\beta_{spacing}$	0.267	41.917	0.006	<.001***
	β_{side}	−0.010	−1.457	0.007	.145
Visual SAC					
	β_{num}	2.045	101.537	0.020	<.001***
	$\beta_{spacing}$	0.270	43.620	0.006	<.001***
	β_{side}	−0.017	−2.568	0.007	.010[a]

Note. β_{num} and $\beta_{spacing}$ are regression weights for the log_2 number and spacing ratios, respectively, and β_{side}reflects a response bias towards one side; *z* is the test statistics of a significance test against zero, *SE* is the standard error of the regression weights.

[a] not significant after Bonferroni adjustment.

necessarily show that number processing is shared between the sensory modalities. The influence of number slightly differed between modalities, giving room for speculation that modality-specific mechanisms might better account for these results. We cannot give a definite answer as to what the exact mechanism is. Nonetheless, there definitely is additional significant influence on the “numerosity” percept, which consequently compromises the strong interpretation of the ANS theory. Even more importantly, we were not able to find a direct correlation between the individual Weber fractions of the haptic and the visual DCT, which would have indicated strong evidence for a modality-shared underlying system processing numerical information. We expected a medium-sized effect between the modalities, which, according to our analysis, is out of the confidence interval boundary. Given that number is the predominant feature involved in the processing of numerosity, the finding that we can reject a medium or stronger relation between haptic and visual DCT performance suggests that number is processed at least partly differently in the visual as compared with the haptic modality.

The ANS theory suggests that numerosity can be extracted from any given stimulus material and is not bound to a modality (Brannon & Merritt, 2011; Dehaene, 2011; Dehaene & Changeux, 1993; Izard et al., 2009; Tokita & Ishiguchi, 2016). In a strong interpretation of this theory, number is an independent feature, and the cognitive system (ANS) is able to filter out all confounding features accompanied with the stimulus material. In conclusion, ANS acuity of participants should be similar across modalities, at least to a high extent. Our results are certainly not consistent with a strong claim of ANS theory. Indeed, we found converging evidence for both the haptic and the visual modality that performance is ratio dependent and that number is the primary source when comparing dot arrays of DCTs. A closer look, however, shows that performance scores were slightly elevated in the haptic condition. This may reflect subtle differences in the DCT implementation of the visual and haptic tasks, or it may indicate factual differences in processing. In addition, the effect of number ratio differed between modalities and even within a modality (i.e., comparing regression coefficients of the visual controls DSC and SAC), and factors beyond number ratio contribute to the responses. Differences between modalities and between visual conditions became visible when we predict the probability of choosing a side (see Fig. 5). For a mutually shared system only extracting numerosity, predictions should have been highly similar (cf. Tokita et al., 2013). A broader interpretation of the ANS would allow for additional factors other than number ratio to play a role. Smaller modality differences in ratio-dependent performance are in principle in line with a broader interpretation of a shared ANS system. However, we can further reject a medium or stronger association between individual ANS acuity in the two modalities, which also contrasts a broader interpretation of the assumption of a single underlying system encoding numerosity in a modality-independent manner.

Our conclusions contradict the interpretations of the conceptual and methodological closest study by Gimbert et al. (2016). The major difference is that we did not accept ratio-dependency as the sole indicator for the involvement of a shared ANS when comparing two different modalities. We think that our enhancements in the haptic stimulus material, and particularly in data analysis, offer supporting evidence for the assumption of a common cognitive system for numerosity encoding between the visual and haptic modality were made precipitously. We think that before numerosity judgments in both modalities can be put into a common framework, substantial differences we found here (i.e., different numerosity weights accompanied by nearly identical spacing weights and lack of cross-sensory associations of individual performance) have to be integrated into a coherent or adapted theory first. In line with results of Clayton et al. (2015), DeWind et al. (2015), and DeWind and Brannon (2016), we suggest that spacing-related factors like convex hull, in addition to number, influence the percept as well. In addition to the already mentioned factors that contribute to the numerosity percept, general cognitive mechanisms, e.g., attention or inhibitory control, may also affect or even explain an association between performances of DCTs (cf. Anobile et al., 2020; Malone et al., 2019). This implies that if an association of performance indices between a visual and a haptic DCT can be found, it will be challenging and important to exactly identify what enabled the comparison between modalities and rule out potential broader mechanisms. Regarding this, we want to emphasize two more aspects, which we think are important to cover in future experiments similar to ours. The first is that participants might (unintentionally) use certain strategies during a task (e.g., use the shape of a pattern as proxy for the numerosity), which can be applied for both modalities and bypass the actual mechanism of interest (i.e., numerosity processing), and therefore compromising results. A way to address this would be to systematically assess participants strategies in the haptic DCT (cf. Dietrich et al., 2019) and generate balanced stimulus patterns (pin patterns), which are less prone to the (effective) use of the most frequent strategies. Another aspect is that participants might respond to a nonnumerical, quantity related dimension of the pins (e.g., surface area). In this case, associations across modalities can be even reliable but reflect quantity comparison rather than numerosity comparison. To examine this further, it could help to clarify the specificity of numerosity estimation in the haptic modality in contrast to a quantity estimation task. This could be achieved with a haptic control task, conceptually similar to a quantity estimation task from Leibovich and Henik (2014), in which participants compare shapes of varying area in a two alternative forced choice task. A resulting performance index in such a task could be used to clarify whether a shared variance to the performance of the numerosity comparison task exists. These, amongst other possible aspects, have to be systematically explored further to improve experiments using visual and especially haptic DCTs.

Overall, we think there is much evidence to shift the perspective from a concept of a shared numerical system that is solely accountable for number estimation to a more integrative system that acknowledges recent findings (Barth et al., 2003; DeWind et al., 2015; DeWind & Brannon, 2016; Gebuis et al., 2016; Leibovich et al., 2017; Smets et al., 2014), rather than to dwell on not having measured the "pure ANS" construct by focusing on methodological details, which occurred in the past. We think that if the ANS is not robust enough to encode numerosities across reasonably similar settings, as, is the case in our study, it is feasible to change the theoretical perspective accordingly.

Different authors were arguing that the ANS theory in its core is not sufficient anymore and needs either revision, for instance a more differentiated approach of quantity processing (Leibovich et al., 2017), or a complete shift in perspective on quantity estimation (Gebuis et al., 2016). For example, Gebuis et al. (2016) argue that ANS theory struggles to explain behavioral as well as neurophysiological data in several instances and a more general system they call "sensory integration system" would better account for reoccurring phenomena regarding confounds. The sensory integration theory differs in two essential points from the ANS theory: (a) sensory cues are integrated into the percept rather than filtered out and (b) the resulting estimation is not an abstract (i.e., directly comparable) number (Gebuis et al., 2016). Following that idea, any "confounding" variable would have weight in the percept and number is estimated by an overall integration of the stimulus features. Taking this theoretical perspective, it would be almost implausible that a "number comparison task" like our DCTs in different modalities and with stimulus features, where only number is kept constant and other potentially equally important features determine the estimate, would result in converging percepts. Taking this into consideration and the fact that number itself seems to be perceived differently between modalities, the lack of correspondence between DCTs in our experiment is comprehensible.

The field of ANS-research generally focuses on group-level comparisons (Halberda et al., 2008; Park & Brannon, 2013; Price et al., 2012), just as we did here. However, this may neglect that individuals differ in their perceptual processes (e.g., perceptual speed), which may be orthogonal to the numerosity comparison acuity. An adjustment for individual processing differences will allow for a better decision whether or not a "truly amodal" ANS exists. In this work, we focused mainly on matching the DCTs between modalities. Although, a part of matching the affordances of the DCTs in both modalities required some specific adjustments which led to task differences (e.g., presentation times, spacing ranges), one may wonder whether this itself can provide an alternative explanation for the absence of performance associations between modalities. Even slight task-variations within a modality can result in significant performance differences as exposed by our visual$_{SAC}$ and visual$_{DSC}$ conditions. Our approach of calculating Weber fractions with the DeWind formula seems favorable to us, as it accounts statistically for implementation differences. Nevertheless, it is important to notice that Weber fractions as performance indices themselves are not a guarantee to have an abstract comparable index for numerosity comparison acuity devoid of context (Guillaume & van Rinsveld, 2018). However, it allowed us to estimate influence of different features to the percept and deviations within a paradigm can at least partly be compensated if sufficient performance indices are used (DeWind & Brannon, 2016), like we did.

Another point is that even though we increased the number of trials in our study to a reasonable degree, it has not been enough to get a good reliability for the individual performance in the haptic DCT. DeWind and Brannon (2016) pointed out that poor correlations are possible because of a lack of reliability in only one dot comparison task (DeWind & Brannon, 2016). This gives room to speculate whether a correlation between modalities might have become visible if the haptic DCT had been more reliable. We acknowledge this possibility to a certain degree. However, the haptic

DCT showed some consistency between sessions and an attenuation projection to account for minor reliability in the cross-modal correlation reveals that a medium sized positive correlation effect can be statistically rejected. We take the liberty of noting that our study did check reliability issues and only hence allows taking lower reliability into consideration, which elsewhere has been ignored at all. Nevertheless, we acknowledge the rather low reliability of the haptic DCT as a shortcoming, which impedes a fully drawn conclusion that a shared numerosity system between vision and haptics is unlikely. A higher reliability of the haptic DCT would have been desirable to dispel doubts from this side. In future work, to overcome shortcomings with reliability, we would recommend an increase in trials. Reliability generally increases with an increase of experimental trials and high trial numbers in DCTs are highly recommended (Dietrich et al., 2015; Lindskog et al., 2013). Further studies should investigate if it is possible to update a haptic DCT to achieve similar reliability as visual DCTs. Another recommendation is to further converge and consolidate task specifications between the DCTs of the visual and the haptic modality. This concerns, for example, overall variety of dot patterns in the haptic modality, which a participant examines within an experiment. Fully randomized dot patterns in every trial (a quite laborious approach) would allow confirmation of performance differences we found between modalities. Another important refinement would be to include and systematically vary the variable area in the haptic DCT to which we were agnostic here. We think that a clean implementation of such a control condition requires further investigation and is even topic for several separate studies because it is possible that a direct transfer from the visual DCT to the haptic DCT setting is not consequently appropriate or applicable in active touch. An additional implementation of these factors in a model like the one in the current study could give deeper insights into the mechanisms that contribute to the numerosity percept in the haptic modality.

Conclusion and outlook

The current study contributes two major findings to the field: Firstly, strong claims of the ANS theory (i.e., number is strictly independent from other features) are not supported by our data. We doubt that neither with preceding reported haptic DCTs nor with our haptic or visual DCT, participants would respond only to number during the task. Secondly, the claim of an amodal shared ANS is also not supported by our data, given that visual and haptic number estimation performances were not related, at least not moderately or more. Following the idea of a sensory integration system, it will be a challenging but important task to disentangle the relevant factors that are integrated into the numerosity percept in the haptic modality. Further research is needed to clarify what the actual similarities are in the visual, haptic, or the auditory modality integration. For now, we think it is feasible to assume that humans can encode numerosity from haptic stimulus material with additional influence of number-related features, which seem to be similar but not equivalent in both the haptic and the visual modality.

Acknowledgments This research was supported by Deutsche Forschungsgemeinschaft (SFB/TRR135/1-2, A05, project number 222641018). Data from the experiment presented here will be available at zenodo.org, https://doi.org/10.5281/zenodo.5763717, starting from publication of the paper.

Conflict of interest statement The authors have no competing interest to declare.

Funding Open Access funding enabled and organized by Projekt DEAL.

References

Anobile, G., Arrighi, R., Castaldi, E., Grassi, E., Pedonese, L., Moscoso, P. A. M., & Burr, D. C. (2018). Spatial but not temporal numerosity thresholds correlate with formal math skills in children. *Developmental Psychology*, *54*(3), 458–473. https://doi.org/10.1037/dev0000448

Anobile, G., Castaldi, E., Moscoso, P. A. M., Burr, D. C., & Arrighi, R. (2020). "Groupitizing": A strategy for numerosity estimation. *Scientific Reports*, *10*(1), 13436. https://doi.org/10.1038/s41598-020-68111-1

Anobile, G., Cicchini, G. M., & Burr, D. C. (2016). Number As a Primary Perceptual Attribute: A Review. *Perception*, *45*(1/2), 5–31. https://doi.org/10.1177/0301006615602599

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition*, *86*(3), 201–221. https://doi.org/10.1016/S0010-0277(02)00178-6

Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(39), 14116–14121. https://doi.org/10.1073/pnas.0505512102

Bertamini, M., Zito, M., Scott-Samuel, N. E., & Hulleman, J. (2016). Spatial clustering and its effect on perceived clustering, numerosity, and dispersion. *Attention, Perception & Psychophysics*, *78*(5), 1460–1471. https://doi.org/10.3758/s13414-016-1100-0

Brannon, E. M., & Merritt, D. J. (2011). Evolutionary foundations of the Approximate Number System. In *Space, time and number in* *the brain* (pp. 207–224). Elsevier. https://doi.org/10.1016/B978-0-12-385948-8.00014-1

Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology, 1904–1920*, *3*(3), 296–322. https://doi.org/10.1111/j.2044-8295.1910.tb00207.x

Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in Cognitive Sciences*, *14*(12), 534–541. https://doi.org/10.1016/j.tics.2010.09.007

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

Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical methods for comparing regression coefficients between models. Advance online publication. https://doi.org/10.1086/230638

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2015). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Routledge.

Craig, J. C., & Lyle, K. B. (2001). A comparison of tactile spatial sensitivity on the palm and fingerpad. *Perception & Psychophysics*, *63*(2), 337–347. https://doi.org/10.3758/bf03194474

Dehaene, S. (2011). *The number sense: How the mind creates mathematics* (Rev. and updated ed.). Oxford University Press. http://gbv.eblib.com/patron/FullRecord.aspx?p=716741

Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience*, *5*(4), 390–407. https://doi.org/10.1162/jocn.1993.5.4.390

DeWind, N. K., Adams, G. K., Platt, M. L., & Brannon, E. M. (2015). Modeling the approximate number system to quantify the contribution of visual stimulus features. *Cognition*, *142*, 247–265. https://doi.org/10.1016/j.cognition.2015.05.016

DeWind, N. K., & Brannon, E. M. (2016). Significant Inter-Test Reliability across Approximate Number System Assessments. *Frontiers in Psychology*, *7*, 310. https://doi.org/10.3389/fpsyg.2016.00310

Dietrich, J. F., Huber, S., & Nuerk, H.-C. (2015). Methodological aspects to be considered when measuring the approximate number system (ANS)—A research review. *Frontiers in Psychology*, *6*, 295. https://doi.org/10.3389/fpsyg.2015.00295

Dietrich, J. F., Nuerk, H.-C., Klein, E., Moeller, K., & Huber, S. (2019). Set size influences the relationship between ANS acuity and math performance: A result of different strategies? *Psychological Research*, *83*(3), 590–612. https://doi.org/10.1007/s00426-017-0907-1

De Smedt, B., Noël, M.-P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, *2*(2), 48–55. https://doi.org/10.1016/j.tine.2013.06.001

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Gebuis, T., Cohen Kadosh, R., & Gevers, W. (2016). Sensory-integration system rather than approximate number system underlies numerosity processing: A critical review. *Acta Psychologica*, *171*, 17–35. https://doi.org/10.1016/j.actpsy.2016.09.003

Geisser, S., & Greenhouse, S. W. (1958). An extension of box's results on the use of the *F* distribution in multivariate analysis. *The Annals of Mathematical Statistics*, *29*(3), 885–891. https://doi.org/10.1214/aoms/1177706545

Gimbert, F., Gentaz, E., Camos, V., & Mazens, K. (2016). Children's Approximate Number System in haptic modality. *Perception*, *45*(1/2), 44–55. https://doi.org/10.1177/0301006615614448

Ginsburg, N., & Pringle, L. (1988). Haptic numerosity perception: Effect of item arrangement. *The American Journal of Psychology*, *101*(1), 131. https://doi.org/10.2307/1422798

Guillaume, M., & van Rinsveld, A. (2018). Comparing numerical comparison tasks: A meta-analysis of the variability of the Weber fraction relative to the generation algorithm. *Frontiers in Psychology*, *9*, 1694. https://doi.org/10.3389/fpsyg.2018.01694

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

Hayward, V. (2011). Is there a 'plenhaptic' function? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1581), 3115–3122. https://doi.org/10.1098/rstb.2011.0150

Hyde, D. C. (2011). Two systems of nonsymbolic numerical cognition. *Frontiers in Human Neuroscience*, *5*, 150. https://doi.org/10.3389/fnhum.2011.00150

Inglis, M., & Gilmore, C. (2014). Indexing the approximate number system. *Acta Psychologica*, *145*, 147–155. https://doi.org/10.1016/j.actpsy.2013.11.009

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(25), 10382–10385. https://doi.org/10.1073/pnas.0812142106

Jordan, N. C., Kaplan, D., Nabors Oláh, L., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development*, *77*(1), 153–175. https://doi.org/10.1111/j.1467-8624.2006.00862.x

Leibovich, T., & Henik, A. (2014). Comparing performance in discrete and continuous comparison tasks. *Quarterly Journal of Experimental Psychology (2006)*, *67*(5), 899–917. https://doi.org/10.1080/17470218.2013.837940

Leibovich, T., Katzin, N., Harel, M., & Henik, A. (2017). From "sense of number" to "sense of magnitude": The role of continuous magnitudes in numerical cognition. *The Behavioral and Brain Sciences*, *40*, e164. https://doi.org/10.1017/S0140525X16000960

Lindskog, M., Winman, A., Juslin, P., & Poom, L. (2013). Measuring acuity of the approximate number system reliably and validly: The evaluation of an adaptive test procedure. *Frontiers in Psychology*, *4*, 510. https://doi.org/10.3389/fpsyg.2013.00510

Malone, S. A., Pritchard, V. E., Heron-Delaney, M., Burgoyne, K., Lervåg, A., & Hulme, C. (2019). The relationship between numerosity discrimination and arithmetic skill reflects the approximate number system and cannot be explained by inhibitory control. *Journal of Experimental Child Psychology*, *184*, 220–231. https://doi.org/10.1016/j.jecp.2019.02.009

McFadden, D. (1977). *Quantitative methods for analyzing travel behaviour of individuals: Some recent developments*. Cowles Foundation Discussion Papers (No. 474). Retrieved from Cowles Foundation for Research in Economics, Yale University website: https://EconPapers.repec.org/RePEc:cwl:cwldpp:474

Mou, Y., & vanMarle, K. (2014). Two core systems of numerical representation in infants. *Developmental Review*, *34*(1), 1–25. https://doi.org/10.1016/j.dr.2013.11.001

Nieder, A. (2016). The neuronal code for number. *Nature Reviews. Neuroscience*, *17*(6), 366–382. https://doi.org/10.1038/nrn.2016.40

Olsson, L., Östergren, R., & Träff, U. (2016). Developmental dyscalculia: A deficit in the approximate number system or an access deficit? *Cognitive Development*, *39*, 154–167. https://doi.org/10.1016/j.cogdev.2016.04.006

Park, J., & Brannon, E. M. (2013). Training the approximate number system improves math proficiency. *Psychological Science*, *24*(10), 2013–2019. https://doi.org/10.1177/0956797613482944

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences*, *14*(12), 542–551. https://doi.org/10.1016/j.tics.2010.09.008

Price, G. R., Palmer, D., Battista, C., & Ansari, D. (2012). Nonsymbolic numerical magnitude comparison: Reliability and validity of different task variants and outcome measures, and their relationship to arithmetic achievement in adults. *Acta Psychologica*, *140*(1), 50–57. https://doi.org/10.1016/j.actpsy.2012.02.008

R Core Team. (2019). R: A language and environment for statistical computing [Computer software]. https://www.R-project.org/

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

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, *15*(1), 72. https://doi.org/10.2307/1412159

Szkudlarek, E., & Brannon, E. M. (2017). Does the approximate number system serve as a foundation for symbolic mathematics? *Language Learning and Development: The Official Journal of the Society for Language Development*, *13*(2), 171–190. https://doi.org/10.1080/15475441.2016.1263573

Szűcs, D., & Myers, T. (2017). A critical analysis of design, facts, bias and inference in the approximate number system training literature: A systematic review. *Trends in Neuroscience and Education*, *6*, 187–203. https://doi.org/10.1016/j.tine.2016.11.002

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

Tokita, M., & Ishiguchi, A. (2016). Precision and Bias in Approximate Numerical Judgment in Auditory, Tactile, and Cross-modal Presentation. *Perception*, *45*(1/2), 56–70. https://doi.org/10.1177/0301006615596888

Tomlinson, R. C., DeWind, N. K., & Brannon, E. M. (2020). Number sense biases children's area judgments. *Cognition*, *204*, 104352. https://doi.org/10.1016/j.cognition.2020.104352

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis [Computer software]. https://ggplot2.tidyverse.org

World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects (2013). *JAMA*, *310*(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053

Attention, Perception, & Psychophysics
https://doi.org/10.3758/s13414-023-02658-9

The role of spatial information in an approximate cross-modal number matching task

Marco Carlo Ziegler[1] · Luisa Karoline Stricker[1] · Knut Drewing[2]

Accepted: 13 January 2023

Abstract

The approximate number system (ANS) is thought to be an innate cognitive system that allows humans to perceive numbers (>4) in a fuzzy manner. One assumption of the ANS is that numerosity is represented amodally due to a mechanism, which filters out nonnumerical information from stimulus material. However, some studies show that nonnumerical information (e.g., spatial parameters) influence the numerosity percept as well. Here, we investigated whether there is a cross-modal transfer of spatial information between the haptic and visual modality in an approximate cross-modal number matching task. We presented different arrays of dowels (haptic stimuli) to 50 undergraduates and asked them to compare haptically perceived numerosity to two visually presented dot arrays. Participants chose which visually presented array matched the numerosity of the haptic stimulus. The distractor varied in number and displayed a random pattern, whereas the matching (target) dot array was either spatially identical or spatially randomized (to the haptic stimulus). We hypothesized that if a "numerosity" percept is based solely on number, neither spatially identical nor spatial congruence between the haptic and the visual target arrays would affect the accuracy in the task. However, results show significant processing advantages for targets with spatially identical patterns and, furthermore, that spatial congruency between haptic source and visual target facilitates performance. Our results show that spatial information was extracted from the haptic stimuli and influenced participants' responses, which challenges the assumption that numerosity is represented in a truly abstract manner by filtering out any other stimulus features.

Keywords Haptics · Multisensory processing · Visual perception

Introduction

The ability to process discrete quantity can be demonstrated in humans, primates, and other animal species (Bisazza & Gatto, 2021; Brannon & Merritt, 2011; Butterworth, 2010; Dehaene, 2011; Feigenson et al., 2004). A cognitive system, the approximate number system (ANS), is thought to be an innate and evolutionary ancient part of number processing, which is shared among humans and other animals (Brannon & Merritt, 2011; Butterworth, 2010; Dehaene, 2011; Nieder, 2016; Spelke & Kinzler, 2007). The ability to quickly estimate numerosity was most likely advantageous for individual fitness (Lourenco & Aulet, 2022). A rapid estimate of numerosity, for example, may have helped to identify which herd has more prey and ultimately deciding which herd to hunt down. Also in modern human life, the ability to accurately estimate numerosity is beneficial every day. It helps one for example to increase the chance of getting a seat in a train during rush hour (or get into the train at all), since accurately estimating the number of people waiting for an arriving train on different train sections allows one to predict the fastest way into the train. The ANS is assumed to enable a fuzzy representation of number without counting (Park & Brannon, 2013). The ANS is reported to process items, objects, or events >4 (Feigenson et al., 2004; Mou & vanMarle, 2014; Olsson et al., 2016; Park & Brannon, 2013; Spelke & Kinzler, 2007).

Three signatures are commonly attributed to the ANS: ratio dependency, nonsymbolic arithmetic operating, and amodality (Brannon & Merritt, 2011). Ratio-dependency is demonstrated when individuals compare which of two numerosities is the larger (e.g., when comparing two dot arrays). The difficulty of the comparison is determined by the ratio of the two numbers rather than their absolute value (e.g., a 10 to 5 comparison is equally difficult to a 20 to 10 comparison). Subsequently, a 5 to 6 comparison is easier than a 7 to 8 comparison even though the absolute

---

✉ Marco Carlo Ziegler
marco.c.ziegler@psychol.uni giessen.de

[1] Department of Psychological Assessment, Justus Liebig University Giessen, Giessen, Germany

[2] Department of Experimental Psychology HapLab, Justus Liebig University Giessen, Giessen, Germany

difference is 1. Nonsymbolic arithmetic operating is a form of calculation with abstract distinct quantity (e.g., the addition of two sets of dot patterns), which can be performed by children even before formal mathematical education (Barth et al., 2006; Gilmore et al., 2010). The third signature, amodality, is a widely assumed axiom based on the idea that the ANS extracts numerosity as an abstract feature from any suitable stimulus material, thereby assuming that the representation of number is amodal (Brannon & Merritt, 2011; Dehaene, 2011; Dehaene & Changeux, 1993; Gebuis et al., 2016; Tokita & Ishiguchi, 2016).

The ANS theory describes a process in different stages of how the percept of numerosity is shaped (Dehaene, 2011; Dehaene & Changeux, 1993; Gebuis et al., 2016). In the first "sensory" stage, the source stimulus is processed by the receiving modality. In vision, this can be a set of dot arrays in varying sizes and positions. In the second stage, the normalizing phase, all stimulus properties are removed and only a standardized signal for each dot remains. In stage three, the accumulation phase, the standardized signals, as well as some degree of error, are summed up into the final percept of numerosity (Gebuis et al., 2016; cf. Tokita et al., 2013). This understanding of the ANS is sometimes referred to as a "direct ANS model" (Qu et al., 2022). The direct ANS model of numerosity estimation often seems to serve as a default theory in the field of numerical cognition (Dehaene, 2011; Halberda et al., 2008; Hyde, 2011; Spelke & Kinzler, 2007). Competing, but to our experience less prevalent in literature, are "indirect models," which assume that a numerosity percept arises from perceiving indirect surrogate cues (e.g., the spatial distribution of stimulus pattern; Qu et al., 2022). Only recently, alternative perspectives or extensions of the direct model have been proposed (Allïk & Tuulmets, 1991; Clarke & Beck, 2021; Gebuis et al., 2016; Gevers et al., 2016; Leibovich, Katzin, Harel, & Henik, 2017a; Lourenco & Aulet, 2022; Walsh, 2003; Zorzi & Testolin, 2018). These perspectives differ in their general view of whether number is a distinct feature and privileged entity in perception or a construct of surrogate perceptional cues (Lourenco & Aulet, 2022; Qu et al., 2022). Furthermore, they distinguish themselves in the question whether the "number sense" is innate (Spelke & Kinzler, 2007), emerges from general, not domain specific abilities, based on exposition and interaction with the environment (Zorzi & Testolin, 2018) or is nonexistent at all (Gebuis et al., 2016). Gebuis et al. (2016), for example, propose that numerosity estimation is the result of integrating different sensory cues, such as distance between stimuli or their convex hull, which only shape a numerosity estimate as required. Other authors argue for a model that considers both number and other magnitudes "holistically" to shape a magnitude percept (Leibovich, Katzin, Harel, & Henik, 2017a). Recently, Lourenco and Aulet (2022) summarized that nonnumerical features such as space and area are sustained throughout processing as magnitude information encoded along with number; the experience of number is an attention-based "reading out" of an integral representation. However, the scope of this proposal is explicitly limited to the visual number processing (Lourenco & Aulet, 2022). This raises the question of whether and how the stream of numerical and nonnumerical information functions throughout sensory and cognitive processing when the source modality is any other than the visual. In summary, some of the proposals and theories have sparked some controversial discourse (Clarke & Beck, 2021; Leibovich, Katzin, Salti, & Henik, 2017b).

Both model types, the direct model and indirect models, can explain phenomena of approximate number processing (e.g., ratio dependency). However, the traditional direct ANS framework conflicts with accumulating evidence, which demonstrates that nonnumerical information (e.g., the spatial area covered by nonsymbolic stimuli or their spatial arrangement) systematically affects the numerosity percept as well (Clayton et al., 2015; DeWind et al., 2015; Gilmore et al., 2016; Hendryckx et al., 2021; Szucs et al., 2013; Tomlinson et al., 2020). DeWind et al. (2015) found that factors that influence the approximate numerosity perception can be formalized into a taxonomy of number, area, and space. They presented a method to statistically quantify these factors and empirically demonstrated that besides number the features space and area additionally contribute to a numerosity percept in a visual dot comparison task (DeWind et al., 2015; DeWind & Brannon, 2016). A consistent finding is that participants use spatial cues (factor space), such as field area, convex hull, or sparsity of a dot array, when comparing two dot arrays (Clayton et al., 2015; Gilmore et al., 2016; Hendryckx et al., 2021). Bertamini et al. (2016) ran an experiment to explicitly identify effects of the spatial arrangement of dot patterns, such as local clustering and occupancy area, on the numerosity and related percepts. They presented participants visual dot arrays in different spatial configurations while keeping the number of dots constant. Indeed, spatial configuration, including local clustering, influenced participants' perception of numerosity. In summary, accumulating evidence suggests that factors other than number alone play a significant role in numerosity estimation tasks, particularly spatial features. As explained above, these findings are in conflict with the prevalent direct ANS theory account, since nonnumerical information should have been removed in the process of creating a numerosity percept. Note that the lack of removing nonnumerical information also questions the claim that number representations are abstract (Brannon & Merritt, 2011), and that an abstract "pure amodal" numerosity representation by the ANS exists. In addition, the amodality assumption is empirically not deeply founded or even contradicted: Only very few studies have investigated numerosity perception in modalities other than the visual or in cross-modal setups (e.g., via tone sequences; Izard et al., 2009), tactile or vibro-tactile stimuli (Tokita & Ishiguchi, 2016; Uluç et al., 2020). Therefore, little is known about similarities or differences of numerosity percepts that are derived from different modalities. More recently, Ziegler and Drewing (2022) used a paradigm of paired presentation of nonsymbolic number stimuli in the haptic and the visual modality comparing participants' performance in the two modalities. They did not find any associations of the individual participant's performance between both intramodal numerosity comparison tasks, which they interpreted as a contradictory result to the amodality assumption. Within the modalities, they found effects of spatial configuration, such as the sparsity of a dot array, affecting participants' numerosity estimate, which additionally questions the abstractness of numerosity representations (Ziegler & Drewing, 2022).

The few cross-modal studies tend to provide counterarguments against abstract amodal representations. Barth et al. (2003) explicitly conducted cross-modal experiments to determine whether the numerosity representation is perceptual or rather abstract. Their assumption was that if a truly abstract representation of (numerical) magnitude exists, there might be little to no cost for cross-modal comparisons relative to unimodal (within-modality, e.g., visual–visual) comparisons in a same–different task. They focused on visual and auditory modalities and different task formats (cf. Dietrich et al., 2015): Participants were exposed to temporally (sequential; visual: flashes, auditory: beeps) or spatially (paired; visual: arrays) presented numerosity stimuli. They determined for each participant which of the two within-modality comparison tasks was more difficult for them (called the "worse unimodal condition") and compared participants' discrimination performance in this condition to their performance in the cross-modal condition. Barth et al. (2003) reported that participants' performance did not significantly differ between the worse unimodal condition and the cross-modal conditions as long as the task format was kept constant. However, participants did perform worse when both modality and task format were crossed (cross-modal and cross-task conditions), which does not seem plausible if true amodal number representation is achieved by filtering out nonnumerical information (cf. Gebuis et al., 2016). Differences in performance between modalities, especially when crossed with presentation format, appear to be a reoccurring pattern across literature (e.g., Anobile et al., 2018). We think, results like those do not fit the theory of an amodal ANS in its "strong" interpretation (i.e., shaping a numerosity percept solely based on number). Tokita et al. (2013) carried out a study investigating the numerosity percepts of individuals with an emphasis on matched task format. They highly standardized a visual and an auditory approximate numerosity comparison task in a sequential paradigm. Participants compared dots or tone sequences in within-modality conditions as well as in a cross-modality condition, all under the same sequential task paradigm. Tokita et al. (2013) found substantial differences in the variability of participants' performance in the visual and auditory modality. Following these results, they argued against the assumption of a modality-independent numerical representation system.

The above evidence against abstract amodal representations from cross-modal studies relies mainly on sequential presentation of numerosity neglecting influences of spatial information. In this work, we investigate cross-modal approximate numerosity perception using spatial dot arrays in the haptic and the visual modality. Spatial (topological) factors (e.g., sparsity, field area, convex hull, density of dot arrays; cf. Clayton et al., 2015; DeWind et al., 2015) have not been extensively evaluated in cross-modal studies. However, spatial factors have been repeatedly found to be a significant factor that affects participants' performance within unimodal tasks (Clayton et al., 2015; DeWind et al., 2015; Gebuis & Reynvoet, 2011; Ziegler & Drewing, 2022). In comparison to surface-area-related factors, that sometimes have been found to be influential (especially in children; cf. Anobile et al., 2018; Tomlinson et al., 2020), spatial factors demonstrate their impact even more consistently. Studies in the visual domain repeatedly have demonstrated that particularly the convex hull of a dot pattern as a spatial factor seems to be informative and a highly relevant contributor for a numerosity estimate (Clayton et al., 2015; DeWind et al., 2015). At the same time, spatial factors such as the convex hull can be well presented and perceived both to the haptic and the visual modality and hence could provide a potent cross-modal influence on numerosity.

The haptic modality allows the construction of stimuli that preserve the spatial information in the task, e.g., in form of massed arrays of stimuli similar to visual arrays. In a cross-modal number matching task, we tested whether spatial information from one modality (haptic) is transferred into another modality (visual) along with the information about numerosity and used in a numerosity task. With conducting a cross-modal study and focusing explicitly on spatial information influences, we try to take the amodality assumption of the ANS to a strong test. We argue that if the classical ANS assumptions would apply—that is, numerosity is (directly) processed by the ANS and the amodal numerosity percept arises due to a removal of nonnumerical cues, spatial information would not affect the performance of participants when comparing numerosity across modalities.

We designed a cross-modal number matching task, in which participants perceive a numerosity haptically via an array of dowels and then ask them to match the extracted numerosity to one of two visually presented dot arrays. Here, we vary if the matching (correct) visually presented dot array is spatially identical to the haptic pattern or a random arrangement of the correct number of dots varying in spatial attributes of the stimulus pattern. The distractor dot array is varied in number, resulting in different ratios, which determine the difficulty of the trial. This allows us to address the following questions:

a) Are individuals able to use *numerical* information extracted from haptically presented source material and match this information to visually presented target stimuli (cross-modal transfer)?

and

b) is *spatial* information extracted along with number, used in the cross-modal numerosity matching task, and thus affects the responses in the cross-modal numerosity matching task?

Addressing the first question is a confirmation that cross-modal magnitude comparison is possible (as in, e.g., Anobile et al., 2018; Barth et al., 2003; Gallace et al., 2007), but with the advancement that rarely investigated modalities are tested in a cross-modal setup with simultaneously presented source numerosity (dot arrays). This provides the necessary fundament for the second question, that additional information besides numerosity is extracted and used by participants. This would contrast the direct ANS model and widens perspective for alternative proposals.

Methods

Participants

We used G*Power (Version 3.1; Faul et al., 2009) to estimate the required sample size of 44 participants to achieve a power of .95 for a medium sized effect (f = .25) in a 2 × 5 repeated-measures analysis of variance (ANOVA). We recruited 50 undergraduates (37 females, mean age M = 23.22 years, SD = 4.48). All participants were healthy and without any impairments or injuries that influenced their touch sensitivity. Participants had normal or corrected-to-normal vision. Forty-three participants were right-handed, and seven were left-handed.

Every participant gave informed written consent to the study prior to the experiment. Consent followed the Declaration of Helsinki (World Medical Association, 2013) and was approved by the local ethics committee of Fachbereich 06 of the Justus-Liebig-University Gießen (LEK-FB 06). Due to the COVID-19 pandemic, protective measures for participants and the experimenter were implemented that complied with the local university guidelines. None of the protective measures compromised the experimental implementation.

Cross-modal number matching task

In the cross-modal number-matching task, participants compared a haptically presented numerosity (dowel array) to two visually presented numerosities (dot arrays). The participant was instructed to decide which of the two visually presented numerosities contained the same number of dots as in the (immediately) prior presented haptic dowel array. We varied whether the visual target stimulus either is a random arrangement of dots (spatially random [SR]) or matches the exact spatial pattern of the haptic stimulus (i.e., being spatially identical [SI]). Furthermore, we varied the numerosity ratio between the target stimulus and the distractor stimulus so that five different levels of difficulty (1.11, 1.14, 1.20, 1.33. 2.00) were implemented.

While the visual target stimulus in the SI condition takes on its spatial features from the haptic stimulus, we utilized the degrees of freedom in stimulus placement in the SR condition to additionally manipulate spatial attributes of the pattern to generate spatially congruent and incongruent trials. As a key metric for spatial influence, we used a compound index, ”spacing,” defined by DeWind et al. (2015) as the product of convex hull and sparsity of a stimulus array. The convex hull can be illustrated as a polygon defined by a subset of elements (dots) containing all elements of the set. Sparsity is defined by the area of convex hull divided by the number of elements in the stimulus array (DeWind & Brannon, 2016). The spatial metric of the haptic stimulus array can be either congruent to the visual target stimulus or to the distractor stimulus (see Experimental Procedure and Data Analysis sections). In a congruent trial, the difference in the variable “spacing” of the haptic stimulus pattern and the target visual stimulus pattern is smaller than the difference in “spacing” of the haptic stimulus pattern and the distractor stimulus pattern. A trial is incongruent if there is a smaller difference in the space metric (= “spacing” compound index) between the distractor visual stimulus pattern and the haptic stimulus pattern than between target visual stimulus and haptic pattern. With this definition, the spatial influence of the pattern can be seen as a facilitating or conflicting factor to the correct response.

Experimental setup

The experiment took place in a quiet, darkened room in the Faculty of Psychology at Justus-Liebig-University Gießen. Participants sat in front of a 22-inch monitor (visible screen: 47.4 cm × 29.6 cm, brightness: 250 cd/m^2, resolution: 1,680 × 1,050, refresh rate: 60 Hz) at 60-cm viewing distance in front of a table. The monitor was placed on a height-adjustable construction that was put on top of the table (see Fig. 1), allowing participants to see the visually presented stimuli at eye level and giving enough space below to comfortably explore the haptic stimuli. We adjusted the screen height

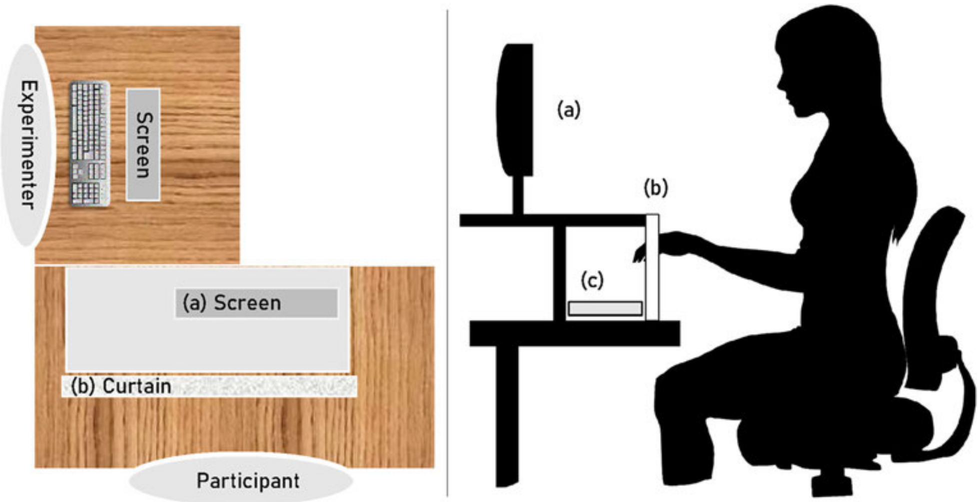


Fig. 1 Illustration of the experimental setup. *Note.* Left image shows a top down view of the setup. Right image shows a cross section view of the setup and how the participants engaged in the task. **a** Participant's screen for visual stimuli; **b** curtain that prevents the participant from visually inspecting the haptic stimulus; **c** haptic stimulus

individually, so that participants looked directly into the center of the screen with upright head posture. The experimenter sat to the left of the participant and could monitor and track the hand during the participant's haptic exploration process. An adjustable opaque piece of fabric (120 × 50 cm) prevented the participant from seeing the haptic stimuli during the experimental trials. Participants could rest their hands under the construction during the experiment. The haptic stimuli were presented 6 cm to the left or 6 cm to the right of the body midline, depending on the participants preferred handedness. The stimulus was locked in place during the trial by a fixture mounted to the table.

Haptic stimuli

The haptic stimuli were wooden panels (length: 119 mm, width: 99 mm, height: 21 mm) made from sanded down veneer plywood ("multiplex"), which could be equipped with varying numbers of wooden dowels. The dowels functioned as single enumerable entity. We prepared 20 slots (diameter: 6 mm, depth: 10 mm) for a panel, which were organized in a 5 × 4 rectangular scheme. The center-to-center distances between the slots were 10 mm, resulting in a 70 mm × 55 mm rectangular area (see Fig. 2). The distance of 10 mm was chosen to enable distinct discrimination of individual dowels in the palm of the hand (Craig & Lyle, 2001). As a reference for the size of the panel and the distances of the slots, we used a glove size S (17.5 cm of hand circumference) to ensure that also individuals with small hand sizes could quickly palpate the stimulus. Any slot could contain an industrial wooden dowel (diameter: 6 mm, length: 30 mm). Each dowel thus protruded 20 mm from the panel. The haptic stimulus could be prepared rapidly during the experiment for the subsequent trial, due to a custom-written software that assisted the experimenter by visualizing the current and next trial. The panels were clipped to the table to ensure stable hold during the exploration phase. In each trial, a panel contained a dowel (dot) array ranging from

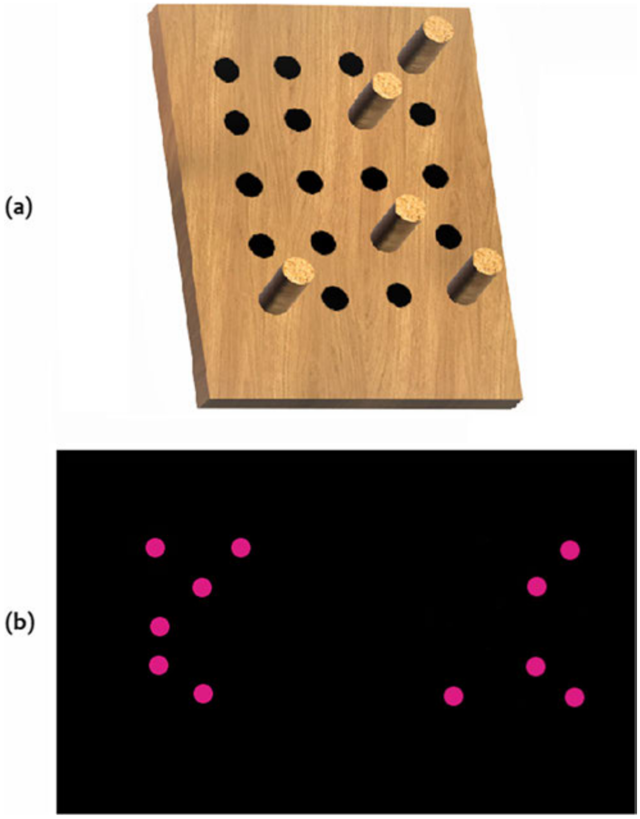


Fig. 2 Examples of haptic and visual stimuli used in the experiment. Note. (***a***) Haptic stimulus. (***b***) Visual stimulus. Haptic pattern matches the visual stimulus on the right side, which displays the identical pattern of dots (SI condition). (Color figure online)

5 up to 10 entities that were randomly (pseudorandomly in the "spatially randomized" condition) selected by the custom-written C++ program. Any position had an equal probability of being selected for the current trial. If a trial did not meet the criteria for the "spatially identical" (SI) condition (i.e., the random pattern was not congruent or incongruent in the spatial attributes), the program recalculated the trial until it matched the criteria (exhaustive search method).

Visual stimuli

Visual stimuli were two simultaneously presented dot arrays on each side of the screen. The arrays were presented 105 mm apart from each other and were clearly distinguishable as separate arrays. A single dot had a diameter of 6 mm on the screen. The minimum center-to-center distance between the dots corresponded with the haptic stimuli and was at least 10 mm on the display. Each array displayed 5 up to 10 magenta-colored dots on black background (see Fig. 2). Possible positions of the dots were arranged in direct correspondence to the possible position of dowels in the wooden panels (5 rows × 4 columns).

Experimental procedure

Each trial started with the participant placing the preferred hand onto the haptic stimulus through the curtain underneath the construction. The experimenter started the trial when the participant touched the stimulus material. A beep tone (duration: 500 ms, pitch: 523 Hz) signaled the participant to start the exploration phase. The participant explored the dowels with the palm of the dominant hand for a total of 4.5 seconds. Participants were prompted to place their hand onto the stimulus several times within the time limit to ensure that they touched all dowels with the palm of their hands, which was monitored by the experimenter. After 3.5 s during the haptic exploration phase, a white fixation cross appeared on the screen. At 4.5 s, a second beep (duration: 100 ms, pitch: 523 Hz) signaled the participant to lift the hand from the haptic stimulus (transition phase). 500 ms later, the fixation cross vanished and two dot arrays were presented for a total of 300 ms, sufficient to give an approximate estimate of the numerical quantity presented. Participants were instructed to choose the side that matched the number of dowels as felt on the panel by verbal announcement "left" or "right," respectively. The response was documented by the experimenter and the next trial began. The full procedure is illustrated in Fig 3.

The experimenter prepared (i.e., equipped the panel with dowels) and exchanged the haptic stimulus in each trial according to a custom-written C++ script. The haptic dowel array was presented to the experimenter by a second screen, which was not visible to the participant. A total of 80 trials were presented. The trial list was shuffled randomly for each participant. Participants were instructed that there is always one visual dot array with the exact same number of dots as dowels and that the other one contained either fewer or more dots.

The brief presentation time of the stimuli (signal to uncover at 4.5 seconds haptic; 300 ms visual) served to prevent the participant from counting the dowels/dots. Pretests were used to determine the appropriate presentation times. Participants were instructed explicitly to base their decision on number. Some participants asked whether the arrangement of the dowels and the arrangement of dots in the dot array might play a role. If this question was posed, the experimenter gave the following answer: "The correct answer is the dot array with the same amount of dots as dowels felt previously. I am not allowed to give any further information about the stimuli." Participants were not given any feedback on whether their answer was correct or not.

In half of the trials, the visual target dot array of correct numerosity was identical (i.e., in numerosity and spatial arrangement) to the dowel array of the haptic stimulus ("spatially identical" [SI]). In the other half of the trials, the correct answer matched only the number of the dowel array, but not the spatial pattern ("spatially randomized" [SR]), which was a spatially random arrangement of dots. Still, the stimuli generating formula allowed balancing the convex hull of the visual stimuli in the "spatially randomized" condition, so that in half of these trials the matching stimulus had a convex hull larger than the distractor stimulus and in the other half a convex hull smaller than the distractor stimulus.

The distractor stimulus in the visual modality was varied in different ratios relative to the haptic stimulus. We applied five different comparison ratios within the experiment, 10:9 (1.11), 8:7 (1.14), 6:5 (1.20), 8:6 (1.33), and 10:5 (2.00). These ratios determine the difficulty if the comparison process is solely based on number. There are easy to distinguish alternatives, for example, haptic stimulus contains 5 dowels, Visual Stimulus 1 shows 5 dots (target), Visual Stimulus 2 shows 10 dots (distractor), and difficult to distinguish alternatives (haptic stimulus contains 9 dowels, Visual Stimulus 1 shows 9 dots (target), Visual Stimulus 2 shows 10 dots (distractor). Each ratio was tested 16 times (8 times in each pattern condition). In 50 % of the trials, the target visual stimulus displayed a lower number of dots than the visual distractor stimulus, and in the other 50 % of trials a higher number of dots. The side on which the "correct" dot array was presented was balanced across trials as well. Two pattern conditions (spatially identical, spatially randomized), 5 ratio conditions, 2 possibilities of whether choosing the higher or lower number as correct target stimulus, 2 convex hull conditions (smaller/larger), and 2 possible sides for the correct answer, resulted in 80 possible combinations ($2 \times 5 \times 2 \times 2 \times 2 = 80$), all of which were tested in a randomized order. The experiment took about 45 minutes to complete.

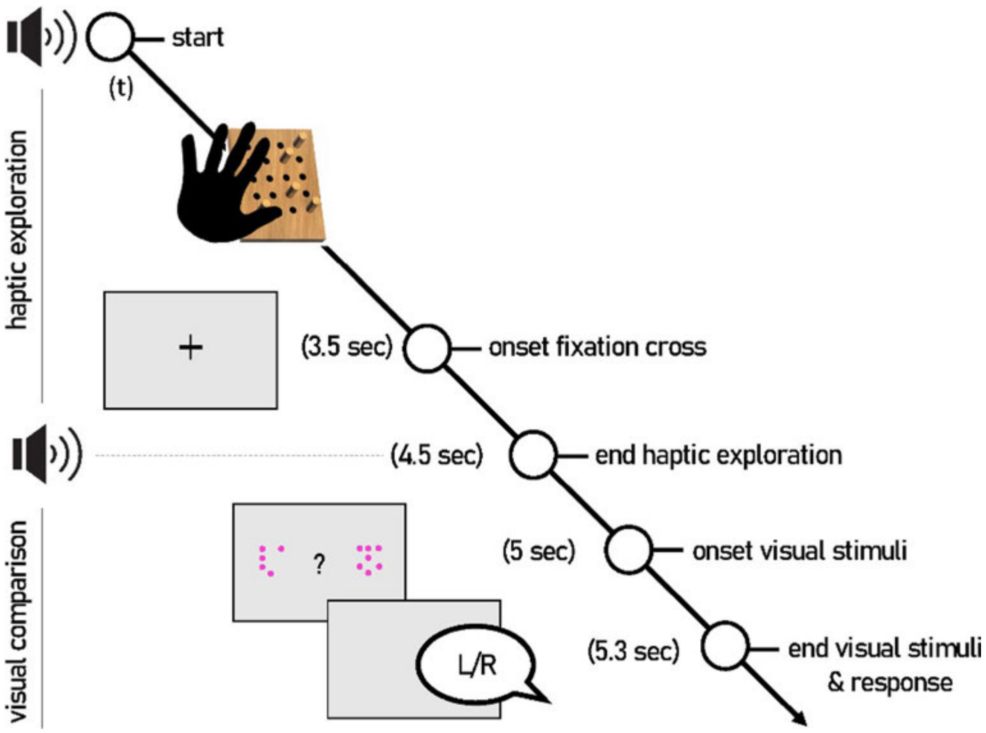


Fig. 3 Illustration of the procedure in the cross modal matching task

Data analysis

For each participant, we calculated the mean of correct answers for each ratio and the two levels of three variables—that is, pattern (SI, SR), magnitude (larger number, smaller number), and bias towards a side (left, right). The factor “magnitude” (larger number, smaller number) allows to check for potential effects on answers in dependence on the relative target magnitude (i.e., whether the target stimulus has the smaller or larger number than the distractor). The bias analysis served as a control condition to detect whether participants had unforeseen tendencies towards a side, regardless of which dot pattern was shown (left side, right side).

We applied arcsine transformation (Cohen et al., 2014) for the aggregated correct answers as preparation for a repeated-measures ANOVA (rmANOVA) with the within-subject variables ratio (2.00, 1.33, 1.20, 1.14, and 1.11) and pattern (spatially identical, spatially randomized). Two additional rmANOVAs with the within-subject variables “ratio” (2.00, 1.33, 1.20, 1.14, and 1.11) and the factors “magnitude” (larger, smaller number) and “bias side” were conducted, respectively. For each rmANOVA, we applied Greenhouse–Geisser correction (Geisser & Greenhouse, 1958) in case of any sphericity violations.

To investigate spatial influence, represented by the product of sparsity and convex hull of a stimulus pattern, on participants’ answers, we analyzed the trials of the “spatially randomized” pattern condition with a generalized linear model (probit link function with iteratively reweighted least squares). The correct answers served as response variable and the variables “ratio,” “spatial congruency” (incongruent/congruent), “magnitude,” and the bias variable (“bias side”) as well as a variable accounting for interindividual differences in performance (“individuals”) as predictors. We use the compound of sparsity and convex hull as an index for spatial influence as it is a relative measure that includes the key spatial features of a stimulus array (DeWind et al., 2015). A trial is spatially congruent if the following condition is met:

$$(Haptic_{space} - VisualTarget_{space}) < (Haptic_{space} - VisualDistractor_{space}) \quad (1)$$

Spatial congruency is therefore a dichotomous variable (congruent, incongruent) reflecting congruency of spatial attributes between the haptic pattern and the visual target pattern.

The data preparation, computation, and visualization of the data were performed using R (Version 4.1.0; R Core Team, 2019).

Results

Table 1 shows the relative frequency of correct trials from all 50 participants sorted by ratio and pattern condition (spatially identical/spatially randomized).

Figure 4 shows the mean percent correct responses of the variable pattern across all applied ratios (2.00, 1.30, 1.20, 1.14, and 1.11).

Table 1 Relative frequencies of correct trials per ratio and pattern in the complete sample

Ratio	Pattern	*N*	*M*	*SE*
1.11	SI	50	.695	.030
1.11	SR	50	.522	.025
1.14	SI	50	.718	.022
1.14	SR	50	.542	.025
1.20	SI	50	.725	.026
1.20	SR	50	.570	.025
1.33	SI	50	.748	.025
1.33	SR	50	.672	.025
2.00	SI	50	.922	.012
2.00	SR	50	.877	.018

Note. "SI" pattern (spatially identical) refers to the condition when the haptic dowel array and the correct visual dot array match in numerosity and also in spatial arrangement. "SR" (spatially randomized) pattern re fers to the condition when the haptic dowel array and the correct visual dot array solely match in numerosity but not in spatial arrangement

Ratio and pattern

The main results of the repeated-measures ANOVA revealed a significant main effect for the factor ratio (2.00, 1.30, 1.20, 1.14, and 1.11) with $F(4, 196)= 59.246, p < .001, \eta_p^2 = .547$ as well as a significant main effect for the factor pattern (spatially identical, spatially randomized) with $F(1, 49) = 59.571, p < .001, \eta_p^2 = .549$. Participants showed a higher accuracy rate with higher ratios. The accuracy rate is higher for spatially identical patterns (see Fig. 4). This trends across all ratios. There is also a significant interaction between ratio and pattern $F(1, 196)= 2.791, p < .001$, but with a smaller effect, $\eta_p^2 = .054$.

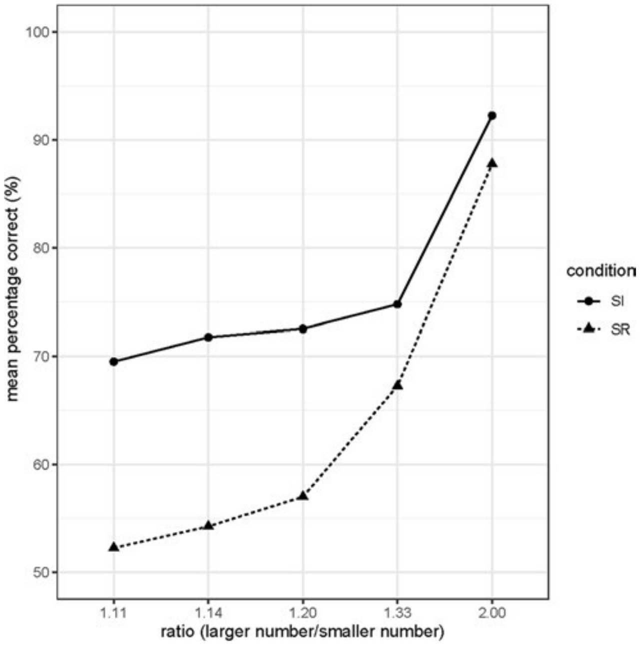


Fig. 4 Mean percentage correct per ratio and pattern in the complete sample. *Note*. "SI" refers to the spatially identical pattern condition when the haptic dowel array and the correct visual dot array match in numerosity and also in spatial arrangement. "SR" refers to the spatially randomized pattern condition, when the haptic dowel array and the correct visual dot array solely match in numerosity but not in spatial arrangement. $N = 50$

Magnitude and bias analysis

We ran repeated-measures ANOVAs to check for effects of number magnitude (larger number, smaller number) and of a potential response bias—that is, a bias towards a side ("bias side," left/right) among the visual stimuli. The bias analysis indicates there was no tendency in responses towards a side, $F(1, 49) = 0.014, p = .906, \eta_p^2 = .000$. Furthermore, there is no interaction effect between the factor "bias side" and "ratio" $F(1, 196) = 1.098, p = .359, \eta_p^2 = .022$. The descriptive statistics are given in Table 2.

The analysis for the factor "magnitude" revealed no significant result, $F(1, 49) = 3.731, p = .059, \eta_p^2 = .071$; as well as no significant interaction between the factor "magnitude" and "ratio" $F(1, 196)= 1.602, p = .175, \eta_p^2 = .032$.

The corresponding descriptive statistics are given in Table 3.

Since the analysis for the variable "magnitude" turned out to be at least marginally significant in the overall sample, $F(1, 49) = 3.731$, $p = .059$, we checked whether this variable is relevant for the subset of trials in the SR condition which is designated for the logistic regression analysis. We found that whether the correct answer was the smaller or the higher number is a significant aspect to consider, $F(1, 49) = 5.448$, $p = .024$. We therefore ran two separate models for the logistic regression, one for trials, in which the higher numerosity was the correct answer and a complementary one, in which the smaller numerosity was the correct answer.

Logistic regression analyses

Within the subset of trials of the "spatially randomized" pattern condition, we distinguished trials where the haptic stimulus was spatial congruent versus spatial incongruent to the visual target stimulus (relative to the relation between haptic stimulus and visual distractor stimulus).

The results of this analysis is given in Table 4. We ran separate models for trials in which the correct target was the higher number (Model A) and trials in which the correct target was the smaller number (Model B). We added the predictors ratio (2.00, 1.30, 1.20, 1.14, and 1.11), spatial congruency (congruent, incongruent) as well as the ID variable ("individuals") representing interindividual differences in participants.

Table 2 Relative frequencies of correct trials per ratio and magnitude in the complete sample

Ratio	Magnitude	*N*	*M*	*SE*
1.11	LN	50	.610	.026
1.11	SN	50	.608	.028
1.14	LN	50	.605	.024
1.14	SN	50	.655	.025
1.20	LN	50	.635	.028
1.20	SN	50	.660	.024
1.33	LN	50	.665	.028
1.33	SN	50	.755	.024
2.00	LN	50	.880	.019
2.00	SN	50	.920	.015

Note. *LN* larger number; *SN* smaller number

Model A

The AIC (Akaike information criterion) of the regression Model A is 1,269.4 with a null deviance of 1,339.3. The estimated pseudo-*R*-squared of the model is at 0.102 (Cragg–Uhler), which is acceptable but gave us reason to ensure the validity of the model fit with the method proposed by Hosmer and Lemesbow (1980). The Hosmer and Lemesbow test does not indicate a poor model fit, $\chi = 6.930$, $df = 8$, $p = .544$. The logistic regression model shows that the number ratios affect the decision for a correct answer the most. Higher ratios (e.g., 2.0) are more likely to provoke a correct answer. Additionally, if a trial showed spatial congruency it facilitated a correct response (see Table 4). The interindividual difference in performance did not show a significant impact within the logistic regression model.

Table 3 Relative frequencies of correct trials per ratio and side in the complete sample

Ratio	Side	*N*	*M*	*SE*
1.11	L	50	.583	.026
1.11	R	50	.635	.029
1.14	L	50	.635	.024
1.14	R	50	.625	.022
1.20	L	50	.650	.024
1.20	R	50	.645	.022
1.33	L	50	.715	.023
1.33	R	50	.705	.023
2.00	L	50	.912	.013
2.00	R	50	.887	.016

Note. "Side" (Left, Right) refers to the response (i.e., left or right side)

Table 4 Model coefficients (Models A and B) and statistics for the logistic regression of correct answers on ratio and spatial congruency

Predictors	β	*Z*	*SE*	*p*
Model A				
Ratio	1.107	7.932	0.139	<.001***
Spatial congruency	0.247	2.997	0.083	<.010**
Individuals	0.003	0.876	0.003	.381
Intercept	1.256	6.102	0.206	<.001***
Model B				
Ratio	1.455	8.905	0.163	<.001***
Spatial congruency	0.111	1 307	0.085	.191
Individuals	0.000	0.065	0.003	.948
Intercept	1.419	6.185	0.229	<.001***

Note. Model A refers to all trials in the "spatially randomized" pattern condition in which the correct target was the higher number; Model B refers to all trials in this condition in which the correct target was the smaller number. "Ratio" includes all ratios 2.00, 1.33, 1.20, 1.14, and 1.11. "Spatial congruency" is a dichotomous variable reflecting congruency in spatial features between the haptic pattern and the visual target numerosity; the predictor "individuals" accounts for interindividual differences in the participants' task performance

$^{**}p \leq .01$. $^{***}p \leq .001$

Model B

The AIC (Akaike information criterion) of the regression model B is 1,181.6 with a null deviance of 1,274.0. The estimated pseudo-*R*-squared of the model is at 0.133 (Cragg–Uhler), which is acceptable as well. The Hosmer and Lemesbow test does not indicate a poor model fit, $\chi = 9.124$, $df = 8$, $p = .332$. The logistic regression model shows that the number ratios affect the decision for a correct answer the most. Higher ratios (e.g., 2.0) are more likely to provoke a correct answer. Spatial congruency did not facilitate a correct response (see Table 4) and was not a significant predictor in Model B. The interindividual difference in performance did not show a significant impact within the logistic regression.

Discussion

In the present study, we examined the ability to compare haptically and visually perceived numerosity from dot arrays in a cross-modal number matching task. We focused on possible effects of spatial information provided by the dot/dowel patterns within this paradigm. We therefore had participants decide which of two visually presented numerosities matched the previously haptically perceived numerosity; we varied whether the target visual stimulus was spatially identical (SI) to the haptic stimulus or whether it displayed a random arrangement (SR) of the same number of dots. The visual distractor stimulus was a numerosity that determined the difficulty of the comparison. The number ratio difficulty ranged from 2.00 (easy) to 1.11 (hard). We furthermore varied whether the visual target stimuli with a spatially randomized pattern (SR) displayed a pattern that is congruent or incongruent to the spatial features of the haptic pattern. We expected that if numerical information is the only processed information in this task, which is implied by the "strong" assumption of the ANS theory, spatial information would not affect the accuracy rates under the variation of the dot patterns (SI/SR). However, our analyses showed a remarkable and significant main effect of the factor "pattern" towards a processing advantage for spatially identical patterns, implying that spatial information is extracted and maintained throughout cognitive processing and has a significant impact on a participant's response. Furthermore, our logistic regression model indicates, at least conditionally, that spatial variables also affect response behavior in the SR condition of our cross-modal approximate number matching task. This demonstrates that participants extracted and used spatial information, besides numerosity, from the given stimulus material even across modalities.

A prerequisite for this study was to confirm that individuals are able to use numerical information extracted from haptically presented source material and match this information to visually presented target stimuli. The accuracy rates in both task conditions show ratio dependency, which is indicative for numerosity processing (Feigenson et al., 2004; Spelke & Kinzler, 2007). Therefore, we can assume that participants compared the number of stimuli, just as instructed, which is consistent with results of previous cross-modal studies (Barth et al., 2003; Gallace et al., 2007; Tokita et al., 2013). What stands out in this study is that in addition to the numerosity of the haptic stimulus, spatial information of the stimulus arrangement was apparently extracted along with number. This result extends previous work by the aspect that spatial information can influence a person's numerosity percept in a cross-modal setup, which has already been shown in intramodal numerosity tasks (Clayton et al., 2015; DeWind et al., 2015; Ziegler & Drewing, 2022). In both conditions, the spatially identical and spatially random pattern, spatial information affected the responses in the cross-modal numerosity matching task. A match in numerosity between the haptic and the visual modality with an identical stimuli arrangement showed the strongest effect and facilitated participants' accuracy significantly. We furthermore demonstrated that correct numerosity discrimination is more likely when the haptic stimulus is congruent in spatial attributes to the visual target stimulus, when the target numerosity was the larger number. The proportions of the effect of ratio and spatial congruency onto the discrimination performance are similar to the effects of numerosity and space found in the haptic-visual intramodal numerosity comparison task by Ziegler and Drewing (2022), which indicates that the same factors contribute to the numerosity percept. Furthermore, the finding of spatial congruency is in accordance with studies that demonstrated intramodal spatial congruency effects (i.e., there is a facilitating effect when spatial and numerical attributes match; Gilmore et al., 2016). Contrary to our expectations, the congruency effect did not show in trials that had the smaller numerosity as target, so the interpretable effect is restricted to the subset of trials in which the target numerosity was the larger one. We assumed that a congruency effect would likewise occur when a smaller "space" goes along with "smaller" numerosity—an assumption we cannot verify in our data.

However, both of these findings, the spatial effect in the SI and the SR condition, are contrary to the assumption that numerosity is achieved by a process that filters out nonnumerical information, which has been proposed by some direct models of approximate numerosity processing (Brannon & Merritt, 2011; Dehaene & Changeux, 1993). Therefore, explanations of the traditional direct ANS theory in how numerosity becomes an amodal representation are challenged as well. The idea of a cognitive system that filters out nonnumerical information to create a pure numerosity percept seems not plausible with our and previous findings.

Nevertheless, it is reasonable to assume that a common representational basis of stimulus information exists in order to compare the magnitudes extracted from different modalities. Our results encourage speculation about the unknown processes and factors involved in numerosity perception and how the results might be embedded in alternative theories. We think key aspects to consider in this discussion and in future work is the extent to which spatial information interacts with numerosity information, whether an integrated numerosity percept is constructed or whether number and spatial information are maintained independently.

Barth et al. (2003) argued for an abstract common magnitude representation that allows comparison between modalities. We partially agree on this, but with the caveat that the numerosity percept is either influenced by or even dependent on spatial factors, or that spatial factors are extracted alongside number and used in the decision stage. The interaction effect between the factor ratio and the pattern condition in our study furthermore implies that the difficulty of a trial, which should be exclusively determined by the ratio, is further modified by the spatial information given. The ratio dependency points towards a mechanism in which spatial information directly shapes the final numerosity percept, which would be in accordance with indirect models of numerosity estimation (cf. Allĭk & Tuulmets, 1991). However, the salient effect of spatial information and consistently better performance across ratios in the SI condition also allows for the plausible interpretation that the numerosity percept and the spatial percept are kept independently but within a shared metric, and that each factor contributes to the decision of choosing the matching "numerosity." In this perspective, the decision whether a numerosity is the same is influenced by these two (and additional) factors, without an irreversible merge into a single numerosity percept. This would follow the interpretation by Lourenco and Aulet (2022) that a shared but distinct representation throughout cognitive processing exists. What remains unclear is to what extent the numerosity factor might be dependent on spatial information assuming that numerosity is derived from topological information, or whether number is a primary feature (cf. Aulet & Lourenco, 2021; Gebuis et al., 2016; Lourenco & Aulet, 2022). In principle, there might be the possibility of an integrated numerosity percept composed of number, space, and presumably area factors, as well as the alternative possibility that space and numerosity are extracted and maintained as distinct multidimensional percepts that can be read out voluntarily, as suggested by Lourenco and Aulet (2022).

Marinova et al. (2021) recently suggested that there might be more than just one way in how numerosity is estimated and that humans are capable of using direct and indirect estimation mechanisms depending on the individual and context. In our context, it would be reasonable to assume that different stimulus material also evokes different information processing within the participant. For example, in our SI condition, participants might have been provoked to emphasize on the spatial information, whereas in the SR condition, due to the lesser spatial salience of the stimulus material, decision making relied more extensively on the numerosity factor of the percept. Still, spatial information remains relevant in the SR condition, but only insofar as the mechanism that derives numerosity from the stimulus pattern is affected by the spatial properties. Additionally, the different effects of spatial congruency hint towards different processing mechanisms within this condition. Ernst and Banks (2002) describe that information integration from different modalities is weighted by its reliability. Cues that are more reliable can "capture" a percept (or decision process). Transferring to our results, spatial features and number features within a task would compete for weighting within the final percept (or at least for the terminal decision, if one assumes that there is no final integration). We think this idea would be plausible and in consonance with the data we present here.

The idea that multiple mechanisms for approximate numerosity estimation exist assumes that there is domain-specific processing of the stimulus material in participants. A limitation of our study might be that participants could have used strategies that bypassed "numerosity" processing and based their decision on other factors, such as a "pattern match" or "shape recognition" of dot arrays. In this view, the results would not directly compromise the traditional direct ANS assumptions, as participants would have just used a strategy to match the task affordances, but cognitive processing could still appear to be rather domain general than domain specific. It is well known that participants supplement their decisions in numerosity comparison tasks with different strategies (Dietrich et al., 2019; Roquet & Lemaire, 2019). A strategy that Roquet and Lemaire (2019) named a "shape"-based strategy could question the striking effect we found for the pattern condition. In this strategy, participants respond to a recognizable shape rather than to specific features like numerosity or space (even though topographic information is still used). Although we cannot completely rule out the possibility that participants occasionally used this strategy, it is very unlikely to explain the effect. Shape recognition was reported as least frequent (0.6 %) in Roquet and Lemaire (2019) in conditions that allowed this strategy to be used even better (intramodal numerosity comparison) than in our setup. However, participants may have selectively used more of what Dietrich et al. (2019) classified as a "numerosity-based strategy" in the "randomized pattern" condition and leaned more toward a "visual strategy" in the "spatially identical pattern" condition. Roquet and Lemaire (2019) used similar terms of "numerical" and "visual" strategies for classification. In our case, visual strategies should be referred to as "topographic" strategies—because the source modality is perceived haptically.

Besides the domain specific processing, other more general aspects in human processing are also involved in a participant's performance, which also might partially explain the performance difference in the SI and SR conditions—for example, the availability of attentional resources. A convincing body of evidence indicates that attentional resources are (at least) partially shared across the visual and haptic modality for tasks that require spatial attention or object-based attention (Wahn & König, 2017). A cross-modal facilitation of performance from one modality to the other due to spatial information could partly be an effect of the shared spatial attention resources between the two modalities (visual, haptic): When participants compared the stimuli (i.e., comparing the haptic percept with the two visual stimuli), a shared spatial pattern of attentional foci between the modalities might have enhanced areas in the spatially identical condition in vision that are already represented from the extraction of the haptic stimulus. That is, visual focal attention in the spatially identical condition might be better tuned to capture the visual stimulus than in the spatially random condition, and thus might have contributed to a better performance in the former condition. Put in other words, in the spatially random condition, the visual and haptic stimuli share less spatial structure in attentional foci as compared to the spatially identical condition and therefore might not enhance performance to the same extent. Our number matching task could have involved both, object-based attention, in which number is the object (analogous to a target letter or a color; cf. Alvarez, 2011), as well as spatial attention, which links to the spatial information of the dot patterns. Both types of attention have recently been demonstrated as influential in visual numerosity estimation tasks (Pomè et al., 2021). Therefore, both could have contributed to the decision process within the cross-modal task, especially in our combination of modalities (visual–haptic). The control of attentional processes could be systematically introduced as factor in experiments with approximate numerosity to better disentangle domain general and domain specific factors that contribute to a participants performance. The role of different types of attention and their effect on performance in numerosity estimation has only recently started to become a greater topic and requires further research (Pomè et al., 2021).

Overall, given the ratio-dependent performance in our task, we have good reason to conclude that participants based their decisions directly or indirectly on numerosity. We furthermore wonder whether strategy use limits any interpretations of the results or whether they are indicative of the distinct processing and if attentional resources partially can explain the performance differences as we have previously speculated.

The (approximate) number of multiple objects is just one feature that can be extracted of a dot pattern nearly instantaneously (Zhang et al., 2019). There is evidence that number is extracted as primary feature (Alvarez, 2011; Feigenson et al., 2004). Additionally, as demonstrated in the visual modality, summary statistics (e.g. averages, variances, orientation, or location of are extracted from objects; Alvarez, 2011). Because selective attention is limited, humans remain being capable to represent multiple aspects of stimuli in these abstracted ways, which are also known as ensemble representations (Alvarez, 2011). These ensemble representations can complement the perception (Alvarez, 2011). Numerosity, per se, seems to be processed distinctly from other statistics in the ensemble representation (Utochkin & Vostrikov, 2017; Yu & Zhao, 2015). However, as mentioned before, a stimulus pattern in our task conveys a variety of information that can be perceived and maintained as ensemble representation throughout cognitive processing. Perceiving sensory information from a haptic stimulus pattern presumably underlies the same principles in the ensemble feature extraction. We wonder, if the spatial information we presented in the SI condition of our task could have been represented and later facilitated the performance of performance just as in the visual modality, through a comparison of the available information in an abstracted summary form (Alvarez, 2011; Alvarez & Oliva, 2009). This would seem plausible to us; however, our experiment did not address this question. Further research will be necessary to clarify the how ensemble statistics extracted from one modality translate into performance of other modalities. We think, that there is an interesting opportunity to find out more about the nature of abstract ensemble representations and if they translate between different modalities and ultimately think all these aspects together in a more general framework of human perception.

In summary, we demonstrate that nonnumerical (i.e., spatial, information is used within our cross-modal number matching task to compare numerosities between the visual and the haptic modalities). We argue that spatial influence affects a person's perception of the given nonsymbolic number stimulus material in our cross-modal matching task, assuming that the final (numerosity) percept includes a composite of number and space (DeWind et al., 2015). We furthermore argue that the so-called approximate number system might be involved in this process, but presumably not according to the strong assumption that this system integrates only numerical information, which has also been questioned in alternative theories and so-called indirect models (cf. Gebuis et al., 2016; cf. Leibovich, Katzin, Harel, & Henik, 2017a). Our data suggests that spatial influence does indeed have a strong impact on participants' decisions, and therefore data interpretation better fits these alternative proposals. We also found that spatial congruency affects the likelihood of accurate numerosity estimations, but it depends on the context: Congruent spatial information facilitated responses when the target number was the larger one, but no congruency effect was observed when the smaller number was paired with smaller spatial magnitude, which should also have been facilitating according to our reasoning. These different congruency effects could be an interesting topic for further research, as congruency effects could help to understand which model of numerosity estimation might be most appropriate (Lourenco & Aulet, 2022).

We articulate concerns about direct ANS conceptions that are narrow in scope and in conflict with recent evidence showing that a numerosity percept is affected by other factors (Clayton et al., 2015; Tomlinson et al., 2020; Ziegler & Drewing, 2022) and therefore, in our opinion, challenge the abstraction claim of the ANS. Albeit speculative, we highlight a range of possibilities for how numerosity processing can work to enable approximate cross-modal number comparison and also allow for flexible adaptation to environmental demands (e.g., via strategy usage). We think that our results also support the suggestion recently proposed by Marinova et al. (2021) that openness to the idea that there may be more than one way in how numerosity is estimated, depending on the individual and context, would better advance this field of research.

Acknowledgments This research was supported by Deutsche Forschungsgemeinschaft (SFB/TRR135, A05, project number 222641018).

Open practice statement Data from the experiment presented here will be available at zenodo.org, 10.5281/zenodo.7117376, starting from publication of the paper.

Funding Information Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The authors have no competing interest to declare.

References

Allïk, J., & Tuulmets, T. (1991). Occupancy model of perceived numerosity. *Perception & Psychophysics, 49*(4), 303 314. https://doi.org/10.3758/bf03205986

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences, 15*(3), 122 131. https://doi.org/10.1016/j.tics.2011.01.003

Alvarez, G. A., & Oliva, A. (2009). Spatial ensemble statistics are efficient codes that can be represented with reduced attention. *Proceedings of the National Academy of Sciences, 106*(18), 7345 7350. https://doi.org/10.1073/pnas.0808981106

Anobile, G., Arrighi, R., Castaldi, E., Grassi, E., Pedonese, L., Moscoso, P. A. M., & Burr, D. C. (2018). Spatial but not temporal numerosity thresholds correlate with formal math skills in children. *Developmental Psychology, 54*(3), 458 473. https://doi.org/10.1037/dev0000448

Aulet, L. S., & Lourenco, S. F. (2021). Perceived number is not abstract. *The Behavioral and Brain Sciences, 44*, Article e179. https://doi.org/10.1017/S0140525X21001102

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition, 86*(3), 201 221. https://doi.org/10.1016/S0010 0277(02)00178 6

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Non symbolic arithmetic in adults and young children. *Cognition, 98*(3), 199 222. https://doi.org/10.1016/j.cognition.2004.09.011

Bertamini, M., Zito, M., Scott Samuel, N. E., & Hulleman, J. (2016). Spatial clustering and its effect on perceived clustering, numerosity, and dispersion. *Attention, Perception, & Psychophysics, 78*(5), 1460 1471. https://doi.org/10.3758/s13414 016 1100 0

Bisazza, A., & Gatto, E. (2021). Continuous versus discrete quantity discrimination in dune snail (Mollusca: Gastropoda) seeking thermal refuges. *Scientific Reports, 11*(1), 3757. https://doi.org/10.1038/s41598 021 82249 6

Brannon, E. M., & Merritt, D. J. (2011). Evolutionary foundations of the approximate number system. In S. Dehaene & E. M. Brannon (Eds.), *Space, time and number in the brain* (pp. 207 224). Academic Press. https://doi.org/10.1016/B978 0 12 385948 8.00014 1

Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in Cognitive Sciences, 14*(12), 534 541. https://doi.org/10.1016/B978 0 12 385948 8.00016 5

Clarke, S., & Beck, J. (2021). The number sense represents (rational) numbers. *The Behavioral and Brain Sciences, 44*, Article e178. https://doi.org/10.1017/S0140525X21000571

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

Cohen, P., Cohen, P., West, S. G., & Aiken, L. S. (2014). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Psychology Press. https://doi.org/10.4324/9781410606266

Craig, J. C., & Lyle, K. B. (2001). A comparison of tactile spatial sensitivity on the palm and fingerpad. *Perception & Psychophysics, 63*(2), 337 347. https://doi.org/10.3758/BF03194474

Dehaene, S. (2011). *The number sense: How the mind creates mathematics* (Rev. and updated ed). Oxford University Press.

Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience, 5*(4), 390 407. https://doi.org/10.1162/jocn.1993.5.4.390

DeWind, N. K., & Brannon, E. M. (2016). Significant Inter Test Reliability across Approximate Number System Assessments. *Frontiers in Psychology, 7*, 310. https://doi.org/10.3389/fpsyg.2016.00310

DeWind, N. K., Adams, G. K., Platt, M. L., & Brannon, E. M. (2015). Modeling the approximate number system to quantify the contribution of visual stimulus features. *Cognition, 142*, 247 265. https://doi.org/10.1016/j.cognition.2015.05.016

Dietrich, J. F., Huber, S., & Nuerk, H. C. (2015). Methodological aspects to be considered when measuring the approximate number system (ANS) A research review. *Frontiers in Psychology, 6*, 295. https://doi.org/10.3389/fpsyg.2015.00295

Dietrich, J. F., Nuerk, H. C., Klein, E., Moeller, K., & Huber, S. (2019). Set size influences the relationship between ANS acuity and math performance: A result of different strategies? *Psychological Research, 83*(3), 590 612. https://doi.org/10.1007/s00426 017 0907 1

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*(6870), 429 433. https://doi.org/10.1038/415429a

Faul, F., Erdfelder, E , Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149 1160. https://doi.org/10.3758/BRM.41.4.1149

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*(7), 307 314. https://doi.org/10.1016/j.tics.2004.05.002

Gallace, A., Tan, H. Z., & Spence, C. (2007). Multisensory numerosity judgments for visual and tactile stimuli. *Perception & Psychophysics, 69*(4), 487 501. https://doi.org/10 3758/BF03193906

Gebuis, T., & Reynvoet, B. (2011). Generating nonsymbolic number stimuli. *Behavior Research Methods, 43*(4), 981 986. https://doi.org/10.3758/s13428 011 0097 5

Gebuis, T., Cohen Kadosh, R., & Gevers, W. (2016). Sensory integration system rather than approximate number system underlies numerosity processing: A critical review. *Acta Psychologica, 171*, 17 35. https://doi.org/10.1016/j.actpsy.2016.09.003

Geisser, S., & Greenhouse, S. W. (1958). An extension of Box's results on the use of the F distribution in multivariate analysis. *The Annals of Mathematical Statistics, 29*(3), 885 891. https://doi.org/10.1214/aoms/1177706545

Gevers, W., Kadosh, R. C , & Gebuis, T. (2016). Sensory integration theory: An alternative to the approximate number system. In A. Henik (Ed.), *Continuous issues in numerical cognition* (pp. 405

418). Academic Press. https://doi.org/10.1016/B978 0 12 801637 4.00018 4

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2010). Non symbolic arithmetic abilities and mathematics achievement in the first year of formal schooling. *Cognition, 115*(3), 394 406. https://doi.org/10.1016/j.cognition.2010.02.002

Gilmore, C., Cragg, L., Hogan, G., & Inglis, M. (2016). Congruency effects in dot comparison tasks: Convex hull is more important than dot area. *Journal of Cognitive Psychology, 28*(8), 923 931. https://doi.org/10.1080/20445911.2016.1221828

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

Hendryckx, C., Guillaume, M., Beuel, A., Van Rinsveld, A., & Content, A. (2021). Mutual influences between numerical and nonnumerical quantities in comparison tasks. *Quarterly Journal of Experimental Psychology, 74*(5), 843 852. https://doi.org/10.1177/1747021820981876

Hosmer, D. W., & Lemesbow, S. (1980). Goodness of fit tests for the multiple logistic regression model. *Communications in Statistics Theory and Methods, 9*(10), 1043 1069. https://doi.org/10.1080/03610928008827941

Hyde, D. C. (2011). Two systems of non symbolic numerical cognition. *Frontiers in Human Neuroscience, 5*, 150. https://doi.org/10.3389/fnhum.2011.00150

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences, 106*(25), 10382 10385. https://doi.org/10.1073/pnas.0812142106

Leibovich, T., Katzin, N., Harel, M., & Henik, A. (2017a). From “sense of number” to “sense of magnitude”: The role of continuous magnitudes in numerical cognition. *Behavioral and Brain Sciences, 40*, e164. https://doi.org/10.1017/S0140525X16000960

Leibovich, T., Katzin, N., Salti, M., & Henik, A. (2017b). Toward an integrative approach to numerical cognition. *Behavioral and Brain Sciences, 40*, e194. https://doi.org/10.1017/S0140525X17000619

Lourenco, S. F., & Aulet, L. S. (2022). A theory of perceptual number encoding. *Psychological Review*. Advance online publication. https://doi.org/10.1037/rev0000380

Marinova, M., Fedele, M., & Reynvoet, B. (2021). Weighted numbers. *The Behavioral and Brain Sciences, 44*, Article e196. https://doi.org/10.1017/S0140525X21001059

Mou, Y., & vanMarle, K. (2014). Two core systems of numerical representation in infants. *Developmental Review, 34*(1), 1 25. https://doi.org/10.1016/j.dr.2013.11.001

Nieder, A. (2016). The neuronal code for number. *Nature Reviews Neuroscience, 17*(6), 366 382. https://doi.org/10.1038/nrn.2016.40

Olsson, L., Östergren, R., & Träff, U. (2016). Developmental dyscalculia: A deficit in the approximate number system or an access deficit? *Cognitive Development, 39*, 154 167. https://doi.org/10.1016/j.cogdev.2016.04.006

Park, J., & Brannon, E. M. (2013). Training the Approximate number system improves math proficiency. *Psychological Science, 24*(10), 2013 2019. https://doi.org/10.1177/0956797613482944

Pomè, A., Thompson, D., Burr, D. C., & Halberda, J. (2021). Location and object based attention enhance number estimation. *Attention, Perception, & Psychophysics, 83*(1), 7 17. https://doi.org/10.3758/s13414 020 02178 w

Qu, C., DeWind, N. K , & Brannon, E. M. (2022). Increasing entropy reduces perceived numerosity throughout the lifespan. *Cognition, 225*, Article 105096. https://doi.org/10.1016/j.cognition.2022.105096

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing https://www.R project.org/

Roquet, A., & Lemaire, P. (2019). Strategy variability in numerosity comparison task: A study in young and older adults. *Open Psychology, 1*(1), 152 167. https://doi.org/10.1515/psych 2018 0011

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science, 10*(1), 89 96. https://doi.org/10.1111/j.1467 7687.2007.00569.x

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 Psychology, 4*, 444. https://doi.org/10.3389/fpsyg.2013.00444

Tokita, M., & Ishiguchi, A. (2016). Precision and bias in approximate numerical judgment in auditory, tactile, and cross modal presentation. *Perception, 45*(1/2), 56 70. https://doi.org/10.1177/0301006615596888

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

Tomlinson, R. C., DeWind, N. K., & Brannon, E. M. (2020). Number sense biases children’s area judgments. *Cognition, 204*, Article 104352. https://doi.org/10.1016/j.cognition.2020.104352

Uluç, I., Velenosi, L. A., Schmidt, T. T., & Blankenburg, F. (2020). Parametric representation of tactile numerosity in working memory. *eNeuro, 7*(1). https://doi.org/10.1523/ENEURO.0090 19.2019

Utochkin, I. S., & Vostrikov, K. O. (2017). The numerosity and mean size of multiple objects are perceived independently and in parallel. *PLoS One, 12*(9), Article e0185452. https://doi.org/10.1371/journal.pone.0185452

Wahn, B., & König, P. (2017). Is attentional resource allocation across sensory modalities task dependent? *Advances in Cognitive Psychology, 13*, 83 96. https://doi.org/10.5709/acp 0209 2

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences, 7*(11), 483 488. https://doi.org/10.1016/j.tics.2003.09.002

World Medical Association. (2013). World medical association declaration of Helsinki: Ethical principles for medical research involving human subjects. *The Journal of the American Medical Association, 310*(20), 2191 2194. https://doi.org/10.1001/jama.2013.281053

Yu, R. Q., & Zhao, J. (2015). Numerosity perception is distinct from mean or sum perception. *Journal of Vision, 15*(12), Article 1030. https://doi.org/10.1167/15.12.1030

Zhang, Y., Liu, T., Chen, C., & Zhou, X. (2019). Visual form perception supports approximate number system acuity and arithmetic fluency. *Learning and Individual Differences, 71*, 1 12. https://doi.org/10.1016/j.lindif.2019.02.008

Ziegler, M. C , & Drewing, K. (2022). Get in touch with numbers An approximate number comparison task in the haptic modality. *Attention, Perception, & Psychophysics, 84*, 943 959. https://doi.org/10.3758/s13414 021 02427 6

Zorzi, M., & Testolin, A. (2018). An emergentist perspective on the origin of number sense. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 373*(1740), Article 20170043. https://doi.org/10.1098/rstb.2017.0043

III Appendix

1 List of all Publications

Ziegler, M. C., Stricker, L. K., & Drewing, K. (2023). The role of spatial information in an approximate cross-modal number matching task. *Attention, Perception, & Psychophysics*. https://doi.org/10.3758/s13414-023-02658-9

Ziegler, M. C., & Drewing, K. (2022). Get in touch with numbers – an approximate number comparison task in the haptic modality. *Attention, Perception, & Psychophysics*. https://doi.org/10.3758/s13414-021-02427-6

Ziegler, M. C., & Zenker, C. A. (2022). *„Which day is tomorrow, if yesterday was the day after Monday?" –can intelligence be measured with "brainteaser" interview questions?*. Poster präsentiert beim 52. Kongress der Deutschen Gesellschaft für Psychologie, Hildesheim, September 2022.

Ziegler, M. C., Stricker, L. K., & Drewing, K. (2022). *How amodal is the "Approximate Number System"? The role of spatial cues in an approximate cross-modal number matching task*. Poster at the 4th conference of the Mathematical Cognition and Learning Society (MCLS), Antwerp, June (2022).

Weingardt, V., Petri, P. S., **Ziegler, M. C.,** & Kersting, M. (2021, Sept.). *Studierendenauswahl: Neues aus der Vergangenheit - Intelligenz und Gewissenhaftigkeit zur Vorhersage von Studienerfolg*. Paper presented at the Biennial Conference of the German Psychological Society Personality Psychology & Psychological Diagnostics (DPPD) Section. 15th September 2021.

Kersting, M., & **Ziegler, M. C.** (2020). Same same, but different. Eignungsdiagnostik auf Distanz. *Personal Magazin, 8*, 34,40.

Ziegler, M. C., & Drewing, K. (2020). *Get in touch with numbers – the relationship of Approximate Number System performance measures in the visual and haptic modality.* Poster at the 3rd conference of the Mathematical Cognition and Learning Society (MCLS), Dublin, June (2020). (Conference canceled)

Ziegler, M. C. (2019). *Possible effects of stimulus range on participants accuracy in Dot-Comparison-Tasks*. Poster präsentiert auf der 15. Arbeitstagung für Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik (DPPD), Dresden, September 2019.

Niemitz, N., Domogala, T., Weingardt, V., **Ziegler, M. C.**, & Kersting, M. (2019). *Difficulty of sample tasks and the threat of instruction texts: Effects on test anxiety and performance*. Poster präsentiert auf der 15. Arbeitstagung für Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik (DPPD), Dresden, September 2019.

Weingardt, V, **Ziegler, M. C.**, Ehrlich S., & Treppesch, C. (2019). *Informationssuche und Nutzung von Online Self-Assessments bei der Studienwahl. Eine vergleichende Untersuchung von Studienanfänger/innen*. Poster präsentiert auf der Fachtagung des Carl-Zeiss-Stiftung Kollegs (CZSK) zur Nutzbarmachung von Online-Self-Assessments für die Studieneingangsphase, Johannes Gutenberg-Universität Mainz, Januar 2019.

Ziegler, M. C., Tammer, K., & Kersting, M. (2018). *Standardisierte Qualitätsminderung- Wie die Praxis der Testinterpretation anhand von Normwerten Objektivität und Validität gefährdet.* Poster präsentiert beim 51. Kongress der Deutschen Gesellschaft für Psychologie, Frankfurt, September 2018.

Ziegler, M. C. (2018). *Denn Sie wissen nicht, wie oft Sie messen sollen! Zur Reliabilitätsmessung des „Approximate Number System" in Abhängigkeit der durchgeführten Experimentaldurchgänge*. Poster präsentiert beim 51. Kongress der Deutschen Gesellschaft für Psychologie, Frankfurt, September 2018.

2 Statement

Statement according to §7, A4 „Promotionsordnung des Fachbereichs Psychologie und Sportwissenschaft der Justus-Liebig-Universität Gießen“ (08.11.2019).

“I hereby declare that I have prepared the thesis at hand independently and without undue aid or the use of any resources other than indicated within the thesis. All parts of my thesis taken either verbatim or analogously from the published or unpublished works of or based on oral communications with others are indicated as such. Regarding all aspects of my scientific enquiries as they appear in my thesis, I have upheld the tenets of good scientific practice as laid out in the "Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis" and complied with the precept of ethics, data protection and animal welfare. I declare that I have neither directly nor indirectly given monetary or any other valuable considerations to others in connection with the thesis at hand. I declare that I have not presented the thesis at hand, either in an identical or similar form, to an examination office or agency in Germany or any other country as part of any examination or degree. All materials from other sources as well as all works performed by others used or directly referenced within the thesis at hand have been indicated as such. In particular, all persons involved directly or indirectly in the development of the thesis at hand have been named. I agree with the screening of my thesis for plagiarism via offline or online detection-software.”

Marco Carlo Ziegler

[redacted]

Marco Carlo Ziegler
(M.Sc.)

Gießen, 08.05.2023