International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: http://www.tandfonline.com/loi/tsed20

Scale and the evolutionarily based approximate number system: an exploratory study

Cesar Delgado, M. Gail Jones, Hye Sun You, Laura Robertson, Katherine Chesnutt & Justin Halberda

To cite this article: Cesar Delgado, M. Gail Jones, Hye Sun You, Laura Robertson, Katherine Chesnutt & Justin Halberda (2017): Scale and the evolutionarily based approximate number system: an exploratory study, International Journal of Science Education, DOI: 10.1080/09500693.2017.1312626

To link to this article: http://dx.doi.org/10.1080/09500693.2017.1312626

Published online: 15 Apr 2017.

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2017
http://dx.doi.org/10.1080/09500693.2017.1312626

Scale and the evolutionarily based approximate number system: an exploratory study

Cesar Delgado[a], M. Gail Jones[a], Hye Sun You[b], Laura Robertson[c], Katherine Chesnutt[a] and Justin Halberda[d]

[a]STEM Education, North Carolina State University, Raleigh, NC, USA; [b]CREATE for STEM Institute, Michigan State University, East Lansing, MI, USA; [c]Curriculum & Instruction, East Tennessee State University, Johnson City, TN, USA; [d]Department of Psychological & Brain Sciences, The Johns Hopkins University, Baltimore, MD, USA

ABSTRACT

Crosscutting concepts such as *scale, proportion, and quantity* are recognised by U.S. science standards as a potential vehicle for students to integrate their scientific and mathematical knowledge; yet, U.S. students and adults trail their international peers in scale and measurement estimation. Culturally based knowledge of scale such as measurement units may be built on evolutionarily-based systems of number such as the approximate number system (ANS), which processes approximate representations of numerical magnitude. ANS is related to mathematical achievement in pre-school and early elementary students, but there is little research on ANS among older students or in science-related areas such as scale. Here, we investigate the relationship between ANS precision in public school U.S. seventh graders and their accuracy estimating the length of standard units of measurement in SI and U.S. customary units. We also explored the relationship between ANS and science and mathematics achievement. Accuracy estimating the metre was positively and significantly related to ANS precision. Mathematics achievement, science achievement, and accuracy estimating other units were not significantly related to ANS. We thus suggest that ANS precision may be related to mathematics understanding beyond arithmetic, beyond the early school years, and to the crosscutting concepts of scale, proportion, and quantity.

ARTICLE HISTORY
Received 19 September 2016
Accepted 24 March 2017

KEYWORDS
Developmental research; cognition; evolutionarily based abilities

National science standards suggest leveraging crosscutting concepts to teach science in a way that not only unites science content, but also integrates other core STEM domains such as engineering and mathematics (NGSS Lead States, 2013). One such crosscutting concept is scale, proportion, and quantity, often called *size and scale* in the literature (NGSS Lead States, 2013). *Size* refers to bulk or quantity, while *scale* refers to the systems of measurement used to compare relative sizes (Magana, Brophy, & Bryan, 2012; Resnick, Davatzes, Newcombe, & Shipley, 2017). In addition to national science standards, scale and measurement are recognised by U.S. mathematics education

CONTACT Cesar Delgado cesar_delgado@ncsu.edu STEM Education, North Carolina State University, Campus Box 7801, Raleigh, NC, USA

standards. ‘Measurement and data’ is one of the major areas emphasised by the *Common core state standards for mathematics* for elementary school mathematics (National Governors Association Center for Best Practices, Council of Chief State School Officers [NGACBP, CCSSO], 2010), and measurement is one of five content strands in the *Principles and standards for school mathematics* (National Council of Teachers of Mathematics [NCTM], 2000). There is a growing need to ensure that students have opportunities for deep learning in science and mathematics to better prepare a scientifically and mathematically literate citizenry capable of making evidence-based decisions in the twenty-first century.

There is a common assumption that a relationship exists between mathematics achievement and science achievement; yet, the degree to which learning science depends on mathematics or is framed by understanding and skills in mathematics has not been fully delineated. The potential for scale to help students integrate their scientific and mathematical knowledge makes investigating scale particularly important as recent international research has found that students and adults in the U.S. trail their international peers in scale and measurement estimation (Delgado, 2013; Jones et al., 2013).

Formal systems of numerical thinking (e.g. mathematics) are culturally derived, yet evidence exists of a relationship between culturally based, complex mathematics and evolutionarily based, innate systems of number (Spelke, 2005). Research in behavioural psychology and neuroscience indicates that mathematical ability is derived from the intersection of two innate core systems of number (Feigenson, Dehaene, & Spelke, 2004a; Feigenson, Libertus, & Halberda, 2013). One core system, referred to as the Exact Number System (ENS) is designated for processing small, exact quantities. The second core system, the approximate number systems (ANS), processes large, imprecise quantities. In addition to behavioural experiments that highlight differences in the ANS and ENS (e.g. Xu, 2003), brain imaging has shown distinct regions of the brain designated for each system (Piazza & Izard, 2009). Furthermore, observations of both the ENS and the ANS in infants and non-human animal species provide evidence that these systems are evolutionarily based rather than culturally acquired (Feigenson et al., 2013; Geary, 1995).

Researchers have argued that these systems are both linked to learning in mathematics. One explanation is that a reciprocal relationship exists between the ENS and the ANS (Libertus, 2015). For example, Halberda and colleagues (Libertus, Odic, Feigenson, & Halberda, 2016) found that strengthening students’ ANS acuity improved performance on mathematics assessments involving exact numerical calculations, and Castronovo and Göbel (2012) found that strengthening the ENS led to improved ANS acuity. The specific mechanisms underlying the relationship between the ENS, the ANS, and learning mathematics and science, however, remain unclear and are in need of further research. A robust field of research has emerged to investigate the relationship that innate systems of number, specifically the ANS, have on learning mathematics in both young children and adults, but less research has been done on older children and even less is known about the relationship between the ANS and concepts in science.

The crosscutting concept of size and scale provides an opportunity to investigate how innate number sense can relate to achievement in science. Size refers to the ability to understand magnitude or quantity of objects, a process that engages the ANS for large numbers (i.e. numbers greater than four). Quantity estimation relies on the ANS to estimate quantity. Similarly, the ANS is also engaged when reasoning about scale because scales are systems of measurement that allow for the comparison of relative sizes, and anytime we map a scale to a number (e.g. 10 inches) the ANS may be involved. Thus, reasoning about concepts of scale may leverage the ANS to discern between two distinct quantities.

Measurement systems and measurement estimation are often used as a means of indexing individuals' understanding of size and scale (Jones & Taylor, 2009). The need for measurement concepts and tools is common for people regardless of location, but the units of measurement that are used vary by culture. In everyday life, the U.S. relies on the U.S. customary (USC) system of units, derived from the English system, while most other countries and scientists worldwide rely on the SI system derived from the metric system (Central Intelligence Agency, 2012). The SI system uses predictable factors across unit prefixes (e.g. there are 1000 nanometres (nm) in a micrometre (μm) and 1000 μm in a millimetre), while the USC system is highly idiosyncratic (e.g. 12 inches in a foot, 3 feet in a yard, and 1760 yards in a mile). Experts tend to create new units that are more convenient and are tied to context (e.g. the light year or astronomical unit) (Jones & Taylor,2009). This process of 'unitizing' underlies conceptual understanding of spatial scales (Tretter, Jones, & Minogue, 2006). Novices also rely on units, spontaneously asking about units smaller than the millimetre when thinking about objects too small to see with the naked eye (Delgado, 2009, 2010). Recent research found significant cross-cultural differences on tasks that assessed accuracy of size estimation across SI-native and USC-native teachers (Jones et al., 2013). Other research with students showed that students who grew up using the SI system for everyday life were more accurate in estimating the length of a metre than their USC-native peers at the same school (Delgado, 2013). There is thus empirical evidence that culturally based knowledge is involved in knowledge of units.

It is possible that our cognition about units of measurement might be built by co-opting the ANS, in which case, there should be a measurable relationship between ANS acuity and accuracy of estimation of the length of standard units. If our knowledge of units is instead built on other biologically primary abilities, such as those that deal with space or with exact numbers, then measures of ANS acuity and knowledge of units may be unrelated. If knowledge of units does not depend strongly on any biologically primary abilities, then measures of ANS acuity and units may also be unrelated.

This study, with participants from a U.S. public school seventh-grade classroom, explored the relationship between ANS acuity and three measures: an assessment of knowledge of standard units of measurement in the SI and USC systems, a state standardised mathematics test, and a state standardised science test.

The goal here is to bring attention to a promising theoretical framework that can inform science education research and practice, and to present findings on whether one evolutionarily based ability is related to concepts of science and mathematic scale at the middle school level. By examining students' knowledge of standard units of measurement, we begin to explore the relationship between the ANS and the crosscutting concept of scale, proportion, and quantity (National Research Council [NRC], 2012). Crosscutting concepts have repeatedly been identified as critical to understanding science and engineering (NGSS Lead States, 2013). Here, we call for more dialogue and research to begin to unravel the complexities inherent in understanding and applying concepts of size and scale.

Theoretical framework

Evolutionarily based abilities

Evolutionarily based abilities have been termed ‘biologically primary’ and are universally acquired even in the absence of external motivation or instruction; they are found in all cultures, plausibly provide an advantage in survival and/or reproduction, and can be traced from related species to humans (Geary, 1995). They include understanding and producing spoken language (Pinker, 1995), and the understanding of space, number, and the behaviour of physical objects and people (Spelke, 1994). These core areas of knowledge relate to essential functions such as interacting with the social and physical environment, navigating the surroundings, and keeping track of dependents, resources, or predators. These areas of knowledge do not depend on individual learning or cultural transmission, as shown by their detection in very young babies; the presence of these abilities or analogous ones in non-human animals suggests evolutionary roots.

There are several evolutionarily based systems of numerical and magnitude representations which may be relevant for scale, proportion, and quantity representation, including the ANS, object tracking, representations of small numbers of individual items, and approximate area and length representations (Cantlon, Platt, & Brannon, 2009; Feigenson, 2007; Feigenson, Dehaene, & Spelke, 2004b; Lourenco, Bonny, Fernandez, & Rao, 2012; McCrink & Wynn, 2004). Among these evolutionarily based abilities, only the ANS has thus far been shown to relate to mathematics learning (Gilmore, McCarthy, & Spelke, 2010; Halberda, Ly, Wilmer, Naiman, & Germine, 2012; Halberda, Mazzocco, & Feigenson, 2008; Libertus, Feigenson, & Halberda, 2011; Libertus, Odic, & Halberda, 2012; Lyons & Beilock, 2011; Mazzocco, Feigenson, & Halberda, 2011a, 2011b), while some recent evidence suggests that area estimation may relate to geometry skill (Bonny & Lourenco, 2015; Lauer & Lourenco, 2016; Lourenco et al., 2012). We have chosen to focus on individual differences in the ANS in the current study, in part because assessment of ANS acuity is a well-developed and well-documented field of study (e.g. Halberda et al., 2008). However, consistent with the suggestion that multiple evolutionarily based and culturally based abilities may be combined when mastering complex reasoning tasks (Paas & Sweller, 2012), we believe that future work might expand to investigate other evolutionarily based abilities. Such research would be greatly aided if assessment tools were developed for measuring individual differences in these other abilities across the lifespan – as has been successfully demonstrated for the ANS.

The ANS, which is the focus of our study, is supported by a neuronal network located in the horizontal segment of the intraparietal sulcus (Cappelletti, Muggleton, & Walsh, 2009; Castelli, Glaser, & Butterworth, 2006; Nieder & Dehaene, 2009). This system processes approximate representations of numerical magnitude allowing humans and other animals to rapidly judge which of two collections of objects is greater in number without explicit verbal counting (e.g. ‘Do we need more chairs to accommodate all of our guests?’), and can allow us to generate noisy (i.e. variable) estimates of numerosity (e.g. ‘About how many marbles are in this jar?’). In the natural environment, the ANS can help animals determine which blackberry bush has more berries, or which herd of prey is larger. The representations of the ANS appear to be functional in infants shortly after birth (Izard, Sann, Spelke, & Streri, 2009). These representations undergo a great deal of refinement and improvement throughout development (Halberda & Feigenson,

2008) and improvements continue throughout the school years with ANS representations attaining their final best precision at approximately 30 years of age (Halberda et al., 2012). This protracted developmental improvement highlights the possible relevance of ANS representations for learning mathematics and other skills.

There is evidence that the ANS is related to mathematics learning. Recent work reveals that ANS acuity at age three years predicts mathematics achievement at age seven (Libertus et al., 2011; Mazzocco et al., 2011a) and that children with a mathematics learning disability have significantly reduced precision in their ANS (Mazzocco, Feigenson, & Halberda, 2011b; Piazza et al., 2010). Neuroimaging studies have shown that the processing of spatial extension (one-dimensional length) and quantity have similar neural bases (Cohen Kadosh et al., 2005; Fias, Lammertyn, Reynvoet, Dupont, & Orban, 2003; Pinel, Piazza, Le Bihan, & Dehaene, 2004). Since humans may 'co-opt' evolutionarily based systems for other uses (Geary, 1995), our understanding of length might be built on the evolutionarily based systems for quantity.

While a link between ANS acuity and quantitative reasoning has been empirically established, the relationship of ANS acuity to estimation of continuous quantities such as length, weight, or volume is not known. There is both evidence suggestive of a link (e.g. from neural data, Pinel et al., 2004) and reasons for caution (e.g. from work on geometry, Lourenco et al., 2012). But, most relatedly, M. Gail Jones (2012) has speculated that there may be a relationship between the ANS and an individual's concepts of size and scale. The present study is to our knowledge the first study investigating the relationship between ANS acuity and concepts of scale at any age, and between ANS acuity and science achievement in middle school.

Culturally based abilities

Culturally based abilities have been termed 'biologically secondary' and are not universally acquired (Geary, 1995). Their development depends on cultural transmission such as formal schooling, and may require external motivation (Genovese, 2003). Abilities such as solving calculus problems, writing, or identifying medicinal plants depend on culturally developed systems of knowledge or tools. Cultural tools are important for thinking and learning (Vygotsky, 1978, 1985). For instance, measurement involves placing identical units across the object to be measured, in a straight line, end to end, without gaps or overlaps, and then summing them; rulers – a cultural tool – do this for us automatically (Lehrer, 2003). As Pea noted, 'intelligence is often distributed by off-loading what could be elaborate and error-prone mental reasoning processes as action constraints of either the physical or symbolic environments' (1988, p. 48). For example, rulers simply do not allow the use of units of different sizes, nor do they permit gaps or overlaps between units.

Differences among cultural tools have been shown to impact the development of knowledge of learners. Some languages, like Mandarin, have very regular number words (eleven is 'ten-one' and twenty-two is 'two-ten-two') while others, like English, have number words that are less transparent in regard to the decimal nature of our numbering system (Delgado, 2013; Miller & Stigler, 1987; NRC, 2009). Research shows that Chinese children can count to higher numbers than their American peers of the same age (Miller & Stigler, 1987). The existence of abilities that emerge at different times in different cultures is a hallmark of culturally based abilities (Geary, 1995). Other cultures have only a few number words (e.g. one, two, and five – Harris, 1987) and combine these to produce other numbers.

Recent international research on teachers' and students' accuracy on scale and measurement tasks showed that there were advantages on some scale and measurement tasks for learners in countries that used the predictable, decimal SI system for everyday life, compared to their American counterparts, who employ the idiosyncratic USC system of units (Delgado, 2013; Jones et al., 2013). Differences in culturally based abilities across cultures have thus been demonstrated for tasks related to the crosscutting concept of scale, proportion, and quantity.

Research objectives

In seeking to explore the relationship between the ANS and scale, proportion, and quantity, the following research questions guide our work:

1) How are scores for mathematics and science achievement related to the evolutionarily based ANS, if at all?
2) How is the accuracy of estimation of various SI and USC measurement units related to the ANS, if at all?

Methods

Participants

The participants were 30 seventh-grade students from a public school located in the south-eastern part of the United States. All participants were 12 years old except for one 11–year-old. The school is a K-12 laboratory school on the campus of a public university, with total enrolment of around 500. Students are selected through a lottery each spring to fill vacant spots. There is one class of students in grades K-6. At the seventh-grade level, a new group of 25 students is added to the class of 25 rising sixth-grade students for a total of 50 seventh graders. Thirteen per cent of the students at the school qualified for free or reduced lunch and the racial/ethnic make-up of the participants is shown in Table 1. Students were all volunteers who were invited to take part in a study of students' concepts of scale and measurement.

The school follows the state curriculum standards for all subjects; the study was conducted prior to the state's adoption of Common Core Standards. The math curricula for grades 6 and 7 at the time of the study included five major categories: mathematical processes; number and operations; algebra; geometry, and measurement; and data analysis, statistics, and probability. The sixth-grade science curriculum covers seven major

Table 1. Demographic characteristics of participants.

	Male	Female	Total
White	6	21	27
Black or African-American	0	2	2
Native Hawaiian or Pacific Islander	1	0	1
Total	7	23	30

topics: inquiry; technology and engineering; interdependence; the universe; the atmosphere; energy; and forces in nature. In seventh-grade science, there are also seven topics: inquiry; technology and engineering; cells; flow of matter and energy; heredity; the earth; and motion. The data were collected during the final week of the school year, so students would have been exposed to the sixth- and seventh-grade curricular topics listed above.

Of 48 students in the seventh grade, 30 participated. Their average scores on the state and math standardised achievement tests were very similar to the overall seventh-grade test scores for the school; by category, their scores varied from 0.3% lower to 3% higher.

Using the publicly available aggregated reports for grades 3–8, the lab school had higher scores on the state standardised tests for science and math than the district, which in turn had higher test scores than the state average. Data for seventh-grade tests also show the school outperforming the state, although not always the district (see Tables 2 and 3). Thus, the setting where this study was conducted was privileged in terms of standardised test scores and had low poverty indicators. Nevertheless, as a pioneering study into ANS and scale, the results are informative, even if not representative of the broader population. Future research should be conducted using a greater diversity of achievement levels, socioeconomic status, and racial/ethnic diversity.

Instruments and variables

We measured students' ANS precision with the test described below, and accuracy in estimating the millimetre (mm), centimetre (cm), metre (m), inch (in), and foot (ft). The most recent mathematics and science achievement scores from the state standardised test were obtained from the school system (these were from the previous school year).

Approximate number system

Students' ANS acuity/precision was assessed with a computer-delivered test that asked respondents individually to quickly judge which of two collections had more objects (http://www.panamath.org/) (Halberda et al., 2008, 2012); see Figure 1. The collections

Table 2. Seventh-grade mathematics standardised test scores for participants, school, district, and state.

	Math processes	Number & operations	Algebra	Geometry & measurement	Data analysis, statistics, & probability
Participants	73	67	53	70	54
School	71	66	52	70	52
District	71	67	53	69	54
State	63	61	47	61	49

Table 3. Seventh-grade science standardised test scores for participants, school, district, and state.

	Cells, flow of matter & energy	Earth	Motion	Inquiry and technology & engineering	Heredity
Participants	73	72	77	71	77
School	72	72	76	70	77
District	64	65	71	67	69
State	61	64	68	65	68

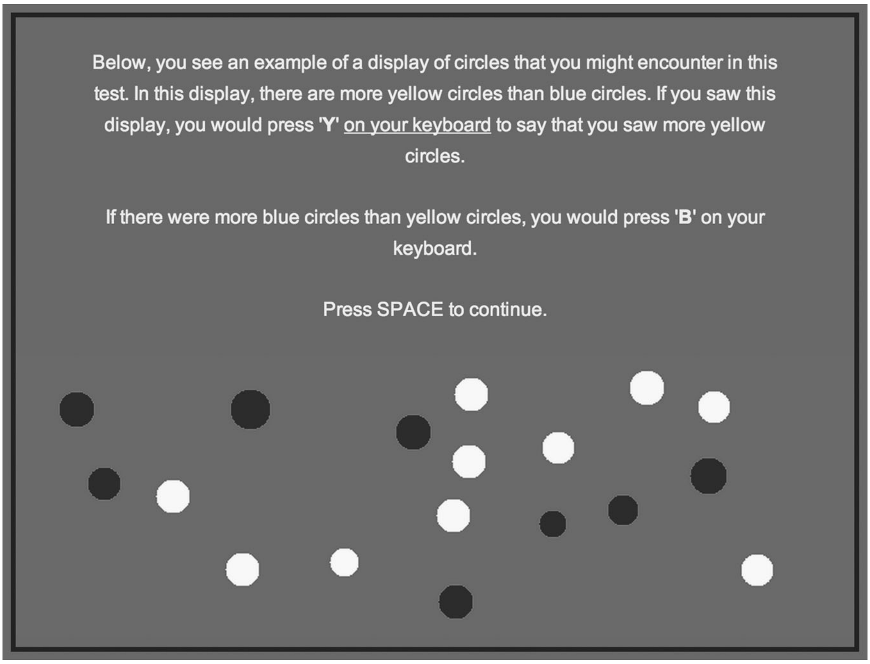

Figure 1. Screenshot of the Panamath computer-delivered test for ANS. The images appear on the screen for 600 milliseconds.

on the test were balanced for other variables that might influence the decision: total area covered by the objects and individual item size. This type of numerical comparison is easy with large ratio differences (e.g. 7 blue dots vs. 14 yellow dots), but becomes more difficult when the ratio of blue to yellow dots is close (e.g. 8 blue vs. 7 yellow) (Dehaene, 1992). By varying the ratio of blue and yellow dots across trials, it is possible to determine the precision of an individual's ANS. The representations of the ANS are thought to be an approximate mental number line where each numerosity is represented by a 'noisy', that is, variable, Gaussian curve centred on the number (Feigenson et al., 2004a). These curves overlap each other such that numerical discrimination can be more or less difficult depending on the ratio and on an individual's internal precision. This precision can be indexed by a Weber fraction (w) and an average response time (RT) on the ANS dots test (Halberda et al., 2008, 2012; Halberda & Feigenson, 2008; Libertus et al., 2011, 2012; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004; Pica, Lemer, Izard, & Dehaene, 2004). The Weber fraction w is an estimate of the internal noise, or confusability, of an individual's ANS number representations, and RT is the amount of time an individual takes to make their decision. Previous research has demonstrated that a person with a more precise ANS will make faster and more accurate number decisions on the ANS dots test, whereas a person with a 'noisier', more variable ANS will perform more poorly and take a longer time to answer, often feeling that they are unsure whether there were more blue or more yellow dots (Halberda et al., 2012). In the present work, we combine estimates of w and RT for each student by z-scoring each of these variables and taking the average z-score (named 'ANS precision') such that lower z-scores correspond to lower w (i.e. higher precision) and lower RT (i.e. faster responding). As z-scored variables have a mean of zero and a standard deviation of 1, negative values of ANS precision indicate better ANS scores than positive values. This procedure also controls for speed–accuracy trade-offs that might exist in performance (Feigenson et al., 2013).

One outlier for w was removed from the analyses. The distribution of the remaining data points was normal according to the Kolmogorov–Smirnov test ($p = .200$). The measure of ANS, ANS precision, was thus interval-level and normally distributed.

Accuracy in estimation of units

We measured the accuracy of student estimates of the length of three everyday SI units (mm, cm, and m) and two USC units (in, ft). Students were provided with a long, thin strip of paper with approximate dimensions 6 cm wide by 2 m long, and asked to mark the endpoint of the five units, starting from a predetermined end. The length of each estimate was then measured in one of two ways. For the smaller estimates, the small end of the strip was placed next to a ruler with millimetre markings, scanned and saved in pdf format. Each estimate was then measured by an undergraduate assistant using Adobe's Measurement Tool (see Figure 2). By expanding the scan (usually to 800%), the lengths of unit estimations on the scans were determined in millimetres with precision nominally down to 0.01 mm. The undergraduate assistant also measured a 10 cm expanse on the ruler, in order to compensate for any deviations from true scale on the scans. We then divided each length of unit estimation measured from the scan by one-hundredth of the 10 cm span measured from the scan in order to arrive at the actual size of the estimation. This was possible for estimations of less than around 25 cm in length, due to the size of the scanner bed. For longer estimations, the fourth author measured them using a ruler.

The per cent error for each measurement unit estimation was then calculated by finding the absolute value of the difference between the estimate and the unit, then

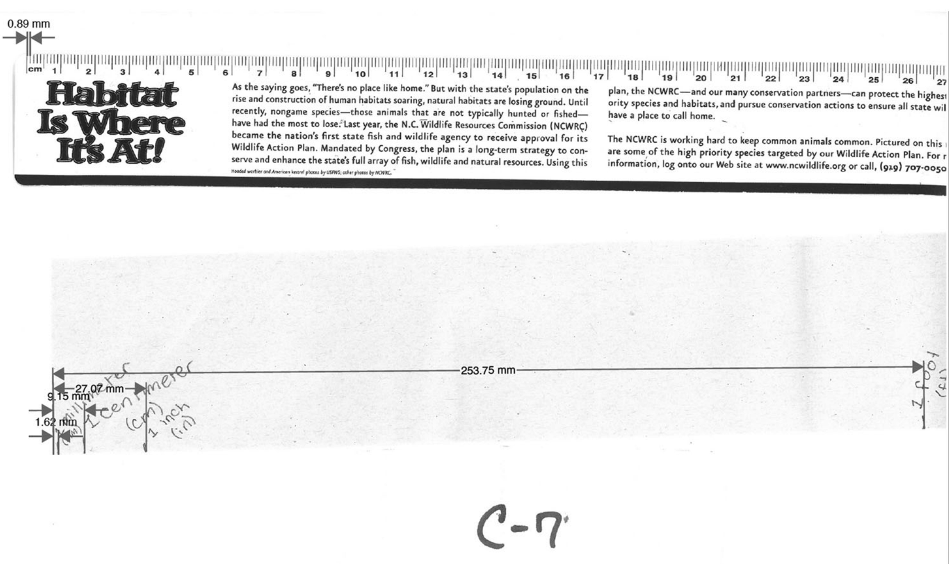

Figure 2. Example measurement of estimation of units using Adobe's Measurement Tool (student C7).

dividing by the unit and multiplying by 100. Accuracy estimating the metre and the foot were normal according to the Kolmogorov–Smirnov test ($p = .076$ and .055, respectively), while accuracy estimating millimetre, centimetre, and inch were non-normal. As described below, the per cent error was used for correlation tests. However, we also recoded these variables into binary variables using a 25% threshold. We used 25% as a threshold in keeping with previous studies (Delgado, 2013; Swan & Jones, 1980). This threshold meant that estimates between 75 and 125 cm in length were considered accurate for the metre estimation task, for example. The resulting variable was dichotomous and ordinal. These dichotomous variables were used in *t*-tests to determine effect size, as described below.

State standardised test scores

The state standardised test scores for mathematics and science were used as a measure of achievement in the two subject areas. Each of the tests was given on one day in the course of a four-day testing schedule of language arts, math, science, and social studies. The assessments were paper-and-pencil tests composed of two parts for each subject area. Each part contained approximately 30 multiple-choice questions and students had approximately 45 minutes to complete it. The mathematics test included five categories of questions: mathematical processes; number and operations; algebra; geometry and measurement; and data analysis, statistics, and probability. Students were provided with a paper ruler and calculator to use during the math test. The science test also included five categories: inquiry and technology & engineering; cells, flow of matter & energy; heredity; the Earth; and motion.

For the participants, most of the categories and overall scores in the state standardised mathematics and science tests were normally distributed according to the Kolmogorov–Smirnov test, while the Science Inquiry and Heredity categories and the Mathematics Algebra category were non-normally distributed.

Data analysis

The relationship between ANS precision, on the one hand, and the state standardised test categories and overall score, on the other, was investigated using Pearson correlations for the normally distributed test scores, and Spearman non-parametric correlations for the non-normally distributed categories.

The relationship between ANS precision and accuracy in estimating each of the five units was investigated in two ways. First, we used correlations to test for the relationship between the variables as a whole, using Pearson correlations for the normally distributed metre and foot and Spearman non-parametric correlations for the non-normally distributed inch, millimetre, and centimetre. Second, we compared the values of ANS precision for the students with accuracy above and below the 25% threshold, using independent-sample *t*-tests for the normally distributed variables and Wilcoxon–Mann–Whitney tests for the non-normally distributed ones. We also calculated effect sizes for statistically significant results, using the pooled standard deviation (i.e. the square root of the average of the squares of the standard deviations of both groups). In order to compensate for multiple comparisons (between ANS and the five units), *p*-values were adjusted using the Bonferroni technique. We used these two analyses in order to be able to compare to previous findings in the literature and calculate effect sizes, using the dichotomous 25% threshold; and to get a better feel for the relationship between the continuous variables, using the correlations. The latter study compensates for the rather arbitrary nature of the 25% threshold, and provides a more robust measure of the relationship between ANS precision and accuracy in measurement estimation.

Results

Research question 1: relationship between ANS precision and state standardised test

None of the correlations between ANS precision and mathematics and science achievement state standardised test scores by categories and overall was statistically significant, for either mathematics or science. The correlations were all negative in value (better ANS precision – represented by lower values – is associated with higher test scores) and weak in magnitude (see Tables 4 and 5).

Research question 2: relationship between ANS precision and accuracy of measurement unit estimation

The accuracy of students' estimations of each measurement unit and the relationship between ANS precision and each measurement unit are reported herein. The SI units were more difficult for students to accurately estimate within a 25% threshold than the USC units. The SI units mm, cm, and m were accurately estimated by 5 out of 30 (17%), 9 out of 30 (30%), and 10 out of 29 (35%) of the students, respectively. In contrast, the USC units inch and foot were accurately estimated by 19 out of 28 (68%) and 17 out of 29 (59)% of the students, respectively.

Only the per cent error in estimation of the metre was significantly correlated to ANS precision. The accuracy in estimating a metre was correlated to ANS precision at a statistically significant level ($p = .007$) and a value of 0.489 for a strong correlation (students

Table 4. Correlations between ANS precision and mathematics standardised test scores ($N = 29$).

	Total Math Score	Math Processes	Number & Operations	Algebra[a]	Geometry & Measurement	Data Analysis, Statistics, & Probability
Correlation	−.227	−.208	−.250	−.125	−.223	−.227
Significance (2-tailed)	.237	.278	.190	.517	.245	.236

[a]Spearman's non-parametric correlation (all others are Pearson correlations)

Table 5. Correlations between ANS precision and science standardised test scores ($N = 29$).

	Total science score	Cells, flow of matter & energy	Earth	Motion	Inquiry and technology & engineering[a]	Heredity[a]
Pearson correlation	−.236	−.162	−.335	−.294	−.196	−.296
Significance (2-tailed)	.217	.401	.076	.122	.309	.119

[a]Spearman's non-parametric correlation (all others are Pearson correlations).

with better ANS precision – represented by a lower number – were better able to accurately estimate the metre – represented by a lower per cent error). Using the Bonferroni procedure to account for multiple comparisons, we compared this *p*-value to the standard *p*-value of .05 divided by the number of comparisons (5), or .01, finding it statistically significant (because $.007 < .01$).

The independent samples *t*-test with the variables for unit estimation dichotomised by the 25% error threshold yielded similar results: the group that accurately estimated the metre had a mean ANS precision of −0.45 while the group with inaccurate estimates had a mean precision of 0.19. The difference is statistically significant with a *p*-value of .005 (with equal variances assumed, as per Levene's test), which is lower than the Bonferroni-adjusted value of .01. The difference in ANS precision between students who were able to estimate the metre accurately and those who were not was 1.246 standard deviation units, for a large effect size. The *t*-test for accuracy estimating the foot and Wilcoxon–Mann–Whitney tests for accuracy estimating the millimetre, centimetre, and inch were not statistically significant.

Discussion

In relation to RQ1, concerning the relationship between ANS and state standardised test scores, the present results contrast with prior research finding statistically significant correlations between ANS and mathematical achievement in early elementary. The trend in the results – for a relationship between basic number sense and more formal measurement knowledge – is in the spirit of the prior research which has found a relationship between basic number sense and more formal math ability – but note that this prior connection for math remains open for continued study both because of the variety of ages yet to be explored and the needed standardisation of the math abilities which participate in this relationship. The lack of statistical significance is likely due in part to low power from the small number of participants. Future research with a larger sample size for greater power should be conducted to further explore the relationship between ANS acuity and achievement in middle school.

In relation to RQ2, it is interesting that there are differences by ANS precision in the estimation of the metre but not the other units, because cultural differences have also shown up in estimating the metre but not other units: students who grew up using the SI system for everyday life were significantly more accurate in estimating the length of a metre than their USC-native peers at the same school but not in estimating a millimetre, centimetre, inch, or foot (Delgado, 2013). These findings must be interpreted with caution because they stem from different studies, yet they are congruent with the idea that both evolutionarily based and culturally based abilities are at play in learning about some units of measurement. This would support Paas and Sweller's proposal that both types of abilities are involved in school learning (2012).

This study used a convenience sample with voluntary participation, which may have limited the generalisability of the results. The lab school setting indicates that our participants had applied to the lottery for admission, and this may indicate a greater interest in education on behalf of the students or their parents, or other systematic differences with typical public school students. In addition, the participants were mainly White, the school has low poverty indicators, and the school has higher state results than the district or state.

Nevertheless, their ANS acuity did not seem to be different from the typical 12-year-old student. While large-scale cross-sectional studies of w and RT have not yet been carried out and there are therefore no standardised estimates of these measures, the mean w for the present group of 12-year-olds ($w = 0.19$) is slightly better than that observed in a sample of 200 10- to 14-year-olds who were part of a large Internet-based study ($w = 0.3$), while the mean RT for the present group (RT = 957 ms) is slightly slower than the Internet sample (RT = 735 ms) (Internet sample published in Halberda et al. (2012). Thus, values for w and RT for the present sample are similar to what has been reported before. The Internet study resulted in slightly faster RT and slightly less accurate responding; since both measures were combined into the ANS precision measure used in this study, it appears that the students in this study were typical of their peers. As for the voluntary nature of participation in the study among the seventh-grade students at the lab school, their state test scores suggest that the participants were not systematically different from their peers. Gender composition was uneven, with 23 female and 7 male participants. However, previous studies have not reported systematic differences in w and RT in girls versus boys. Large cross-sectional studies would be valuable to further investigate these issues.

Implications and future directions

Theoretical implications

This study contributes to the debate on the nature of knowledge that has been ongoing for a decade. It has been suggested that, 'many skills may consist of a combination of primary and secondary knowledge and so we may be dealing more with a continuum than a dichotomy' (Paas & Sweller, 2012, p. 40). The findings reported here, while exploratory, align with this theory, as the accuracy in estimating the length of a metre correlates to ANS precision, while previous research found that it varies across cultural groups (Delgado, 2013; Jones et al., 2013). One could speculate that skills such as estimation of length that have strong relationships to critical human activities and historically would have been essential to survival, would be related to primary, evolutionary-based knowledge.

This study also adds to the research base on evolutionarily based knowledge by showing that middle school standardised test scores for mathematics do not correlate with ANS precision at a statistically significant level, although trends are in the same direction. It also showed that the estimation of some SI units correlates with ANS precision – the first finding for the ANS' possible influence on learning that pertains to science.

Educational implications

Currently, math and science education standards recommend measurement activities using non-standard units before introducing standard units of length starting in the second grade (NCTM, 2000; NGACBP, CCSSO, 2010) or mid-elementary (NRC, 2012). Given that accuracy in estimating some standard units (the metre) is correlated to a biologically primary ability (ANS precision), it is worth exploring whether the introduction of standard units can begin earlier, perhaps using discovery methods or activities that are specifically aimed at mobilising the ANS. For instance, learning to visualise a line or length as a collection of equal-sized components in order to estimate its size or to compare it to another line might rely on the ANS; if those components are standard rather than invented units, then the activities aimed at building the culturally based knowledge of measurement units and length would be more explicitly based on the evolutionarily based ANS.

Future research

Additional studies are needed that can unravel which abilities are biologically based and which emerge in cultural contexts. But even more important is understanding the conditions under which an individual uses biologically primary abilities in conjunction with secondary abilities to accomplish learning tasks. Other areas that have yet to be explored in depth include the intersection of sets of evolutionarily based abilities. For example, spatial visualisation is likely involved in the recall of objects in terms of body lengths, as used in previous research by the authors (Delgado, 2013; Tretter, Jones, & Minogue, 2006; Tretter, Jones, Andre, Negishi, & Minogue, 2006). Spatial visualisation, according to Geary (1995), may have evolved as part of the development of navigation skills. Furthermore, Geary has suggested that spatial visualisation is co-opted and used as part of algebraic problem-solving.

The goal of this paper has been to bring the theoretical framework of biologically primary and secondary abilities to the attention of the science education community, and to present findings that are consistent with a mechanism where both types of knowledge play a role in school learning. We invite other researchers to explore whether the topics and concepts they investigate build on these two different types of ability, and to design studies to further examine how reasoning about the crosscutting concept of size and scale develops, considering that it may involve the co-opting of multiple primary systems and a rich interplay with secondary abilities.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Bonny, J. W., & Lourenco, S. F. (2015). Individual differences in children's approximations of area correlate with competence in basic geometry. *Learning and Individual Differences*, *44*, 16–24.

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences*, *13*(2), 83–91.

Cappelletti, M., Muggleton, N., & Walsh, V. (2009). Quantity without numbers and numbers without quantity in the parietal cortex. *Neuroimage*, *46*(2), 522–529.

Castelli, F., Glaser, D. E., & Butterworth, B. (2006). Discrete and analogue quantity processing in the parietal lobe: A functional MRI study. *Proceedings of the National Academy of Sciences*, *103*(12), 4693–4698.

Castronovo, J., & Göbel, S. M. (2012). Impact of high mathematics education on the number sense. *PLoS ONE*, *7*(4), 1–16. doi:10.1371/journal.pone.0033832

Central Intelligence Agency. (2012). *The world factbook*. Retrieved November 9, 2012, from https://www.cia.gov/library/publications/the-world-factbook/

Cohen Kadosh, R., Henik, A., Rubinsten, O., Mohr, H., Dori, H., Van de Ven, V., … Linden, D. E. (2005). Are numbers special? The comparison systems of the human brain investigated by fMRI. *Neuropsychologia*, *43*(9), 1238–1248. doi:10.1016/j.neuropsychologia.2004.12.017

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

Delgado, C. (2009). *Development of a research-based learning progression for middle school through undergraduate students' conceptual understanding of size and scale* (*Unpublished doctoral dissertation*). University of Michigan, Ann Arbor.

Delgado, C. (2010). Units of length: A notational system for conceptual understanding of size and scale. In *Proceedings of the 9th international conference of the learning sciences – Volume 2* (pp. 362–363). Chicago, IL: International Society of the Learning Sciences. Retrieved from http://portal.acm.org/citation.cfm?id=1854509.1854698

Delgado, C. (2013). Cross-cultural study of understanding of scale and measurement: Does the everyday use of US customary units disadvantage US students?. *International Journal of Science Education*, *35*(8), 1277–1298.

Feigenson, L., Dehaene, S., & Spelke, E. S. (2004a). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314.

Feigenson, L., Dehaene, S., & Spelke, E. S. (2004b). Origins and endpoints of the core systems of number. Reply to Fias and Verguts. *Trends in Cognitive Sciences*, *8*(10), 448–449.

Feigenson, L. (2007). The equality of quantity. *Trends in Cognitive Sciences*, *11*(5), 185–187.

Feigenson, L., Libertus, M. E., & Halberda, J. (2013). Links between the intuitive sense of number and formal mathematics ability. *Child Development Perspectives*, *7*(2), 74–79.

Fias, W., Lammertyn, J., Reynvoet, B., Dupont, P., & Orban, G. A. (2003). Parietal representation of symbolic and nonsymbolic magnitude. *Journal of Cognitive Neuroscience*, *15*(1), 47–56.

Geary, D. C. (1995). Reflections of evolution and culture in children's cognition: Implications for mathematical development and instruction.. *American Psychologist*, *50*(1), 24–37.

Genovese, J. E. C. (2003). Piaget, pedagogy, and evolutionary psychology. *Evolutionary Psychology*, *1*, 127–137.

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. doi:10.1016/j.cognition.2010.02.002

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the 'number sense': The approximate number system in 3-, 4-, 5-, 6-year-olds and adults. *Developmental Psychology*, *44*(5), 1457–1465.

Halberda, J., Ly, R., Wilmer, J., Naiman, D., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive internet-based sample. *Proceedings of the National Academy of Sciences*, *109*(28), 11116–11120.

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

Harris, J. (1987). Australian Aboriginal and islander mathematics. *Australian Aboriginal Studies*, *2*, 29–37. Retrieved June 9, 2014, from http://search.informit.com.au/documentSummary;dn=189499873710770;res=IELIN

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.

Jones, M. G. (2012, May). Invited paper presented at the International STEM Research Symposium: Quantitative Reasoning in Mathematics and Science Education, Savannah, GA.

Jones, M. G., Paechter, M., Gardner, G., Yen, I., Taylor, A., & Tretter, T. (2013). Teachers' concepts of spatial scale: An international comparison. *International Journal of Science Education*, *35*(14), 2462–2482.

Jones, M. G., & Taylor, A. (2009). Developing a sense of scale: Looking backward. *Journal of Research in Science Teaching*, *46*(4), 460–475.

Lauer, J. E., & Lourenco, S. F. (2016). Spatial processing in infancy predicts both spatial and mathematical aptitude in childhood. *Psychological Science*, *27*(10), 1291–1298.

Lehrer, R. (2003). Developing understanding of measurement. In J. Kilpatrick, W. G. Martin, & D. Schrifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 179–192). Reston, VA: National Council of Teachers of Mathematics.

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. doi:10.1111/j.1467-7687.2011.01080.x

Libertus, M. E. (2015). The role of intuitive approximation skills for school math abilities. *Mind, Brain, and Education*, *9*(2), 112–120.

Libertus, M. E., Odic, D., Feigenson, L., & Halberda, J. (2016). The precision of mapping between number words and the approximate number system predicts children's formal math abilities. *Journal of Experimental Child Psychology*, *150*, 207–226.

Libertus, M. E., Odic, D., & Halberda, J. (2012). Intuitive sense of number correlates with math scores on college-entrance examination. *Acta Psychologica*, *141*(3), 373–379.

Lourenco, S. F., Bonny, J. W., Fernandez, E. P., & Rao, S. (2012). Nonsymbolic number and cumulative area representations contribute shared and unique variance to symbolic math competence. *Proceedings of the National Academy of Sciences*, *109*(46), 18737–18742.

Lyons, I. M., & Beilock, S. L. (2011). Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition*, *121*(2), 256–261.

Magana, A. J., Brophy, S. P., & Bryan, L. A. (2012). An integrated knowledge framework to characterize and scaffold size and scale cognition (FS2C). *International Journal of Science Education*, *34*(14), 2181–2203.

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011a). Preschoolers' precision of the approximate number system predicts later school mathematics performance. *PLoS One*, *6*(9). doi:10.1371/journal.pone.0023749

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011b). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). *Child Development*, *82*(4), 1224–1237. doi:10.1111/j.1467-8624.2011.01608.x

McCrink, K., & Wynn, K. (2004). Large-number addition and subtraction by 9-month-old infants. *Psychological Science*, *15*(11), 776–781.

Miller, K. F., & Stigler, J. W. (1987). Counting in Chinese: Cultural variation in a basic cognitive skill. *Cognitive Development*, *2*(3), 279–305. doi:10.1016/S0885-2014(87)90091-8

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. Washington, DC: Author.

National Research Council. (2009). *Mathematics learning in early childhood: Paths toward excellence and equity* (pp. 95–120). Washington, DC: The National Academies Press.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: The National Academies Press.

Nieder, A., & Dehaene, S. (2009). Representation of number in the brain. *Annual Review of Neuroscience*, *32*, 185–208. doi:10.1146/annurev.neuro.051508.135550

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review*, *24*(1), 27–45. doi:10.1007/s10648-011-9179-2

Pea, R. D. (1988). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 47–87). Cambridge: Cambridge University Press.

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., … Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition*, *116*(1), 33–41.

Piazza, M., & Izard, V. (2009). How humans count: Numerosity and the parietal cortex. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *15*(3), 261–273. doi:10.1177/1073858409333073

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*, *44*(3), 547–555.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*(5695), 499–503.

Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, *41*(6), 983–993.

Pinker, S. (1995). *The language instinct* (1st Harper Perennial ed.). New York, NY: Harper Perennial.

Resnick, I., Davatzes, A., Newcombe, N. S., & Shipley, T. F. (2017). Using relational reasoning to learn about scientific phenomena at unfamiliar scales. *Educational Psychology Review*, *29*, 11–25. doi:10.1007/s10648-016-9371-5

Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science?: A critical review. *American Psychologist*, *60*(9), 950–958. doi:10.1037/0003-066X.60.9.950

Spelke, E. (1994). Initial knowledge: Six suggestions. *Cognition*, *50*(1–3), 431–445.

Swan, M., & Jones, O. E. (1980). Comparison of students' percepts of distance, weight, height, area, and temperature. *Science Education*, *64*(3), 297–307. doi:10.1002/sce.3730640305

Tretter, T. R., Jones, M. G., Andre, T., Negishi, A., & Minogue, J. (2006). Conceptual boundaries and distances: Students' and experts' concepts of the scale of scientific phenomena. *Journal of Research in Science Teaching*, *43*(10), 282–319.

Tretter, T. R., Jones, M. G., & Minogue, J. (2006). Accuracy of scale conceptions in science: Mental maneuverings across many orders of spatial magnitude. *Journal of Research in Science Teaching*, *43*(10), 1061–1085.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1985). *Thought and language*. (A. Kozulin, Ed.). Cambridge: MIT Press.

Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition*, *89*(1), B15–B25.