How the abstract becomes concrete:

Irrational numbers are understood relative to natural numbers and perfect squares

Purav Patel[1], Sashank Varma[2]

[1]Department of Educational Psychology, University of Wisconsin – Madison, Madison, WI, USA

[2]Department of Educational Psychology, University of Minnesota – Twin Cities, Minneapolis, MN, USA

Corresponding Author:

Purav Patel

Room 859, Educational Sciences Building
Department of Educational Psychology
University of Wisconsin-Madison
1025 West Johnson Street
Madison, WI 53706-1796

Email: patel011393@gmail.com

Keywords: irrational, perfect square, magnitude, number line, strategy, arithmetic, reference, individual differences

Abstract

Mathematical cognition research has largely emphasized concepts that can be directly perceived or grounded in visuospatial referents. These include concrete number systems like natural numbers, integers, and rational numbers. Here, we investigate how a more abstract number system, the irrationals denoted by radical expressions like $\sqrt{2}$, is understood across three tasks. Performance on a magnitude comparison task suggests that people interpret irrational numbers – specifically, the radicands of radical expressions – as natural numbers. Strategy self-reports during a number line estimation task reveal that the spatial locations of irrationals are determined by referencing neighboring perfect squares. Finally, perfect squares facilitate the evaluation of arithmetic expressions. These converging results align with a constellation of related phenomena spanning tasks and number systems of varying complexity. Accordingly, we propose that the task-specific recruitment of more concrete representations to make sense of more abstract concepts (*referential processing*) is an important mechanism for teaching and learning mathematics.

Introduction

The historical development of number systems can be characterized as a gradual progression from the concrete to the abstract. Natural numbers, which denote the cardinalities of sets, were understood directly and formalized first. They are concrete in the sense that they map directly to quantities in the material world such as the number of berries in one's hand. According to the mathematician Leopold Kronecker, "God created" this relatively concrete number system, whereas more abstract number systems are "the work of man" (Bell, 1986). In other words, more abstract number systems were discovered later and constructed hierarchically upon the natural numbers (Landau, 1966). These systems are still relatively concrete in the sense that they can be interpreted as corresponding to real-world quantities like basement floors for integers and pie pieces for rationals (Van de Walle, Karp, & Bay-Williams, 2010).

It is with the irrational numbers, which include $\sqrt{2}$ and π, that mathematicians discovered a number system lacking material referents or models that build on intuition (Struik, 1987). Such abstraction is associated with many surprising properties. For instance, there are more irrational numbers than natural numbers, integers, or rational numbers. The set of all irrationals is *uncountably* infinite in cardinality, whereas the latter sets are each *countably* finite (Struik, 1987). Mathematics teachers, in an attempt to facilitate understanding, encourage students to think of irrational numbers like π by using rational number approximations like 3.14. Unfortunately, this can lead to conceptual confusions (González-Martín, Giraldo, & Souto, 2013; Sirotic & Zazkis, 2007).

Cognitive science research has left irrational numbers unexamined, focusing instead on more concrete number systems. For example, the natural numbers are thought to be grounded in directly perceptible objects and internally represented as continuous magnitudes on a mental number line (Moyer & Landauer, 1967). For multiple number systems, researchers have acknowledged the necessity of mastering associated symbol systems. However, the focus is often on how these symbol systems recruit and restructure magnitude representations. Examples include incorporating the inverse relationship between positive and negative integers into the mental number line and capturing the ratio relationship between the numerators and denominators of fractions more precisely (Matthews & Chesney, 2015; Varma & Schwartz, 2011).

For more abstract number systems, do magnitude representations continue to be heavily recruited or does notation-specific strategic processing become more important (Bender, Schlimm, & Beller, 2015; Zhang & Norman, 1995)? We addressed this question using irrational numbers as a test case, specifically those denoted by radical expressions like $\sqrt{2}$. We investigated the task demands and stimulus properties that cause adults to rely on magnitude representations versus strategic processing. Our findings reveal the vital roles that concrete natural number referents play in understanding abstract irrational number expressions. They have implications for extending theories of numerical cognition and improving instructional design in mathematics education. Our approach is grounded in prior studies of how people understand natural numbers, integers, and rational numbers.

Natural numbers

The natural numbers are the set of countable numbers $\{0, 1, 2, \ldots\}$.[1] Their mental representation has been studied extensively using the *magnitude comparison* (MC) task. The

---

[1] Mathematicians disagree on whether the natural numbers begin with 0 or 1. The original Peano-Dedekind axioms begin with 1 (Dedekind, 1963/1888). However, more recent presentations of

primary finding is that when comparing which number in a pair is greater or lesser, response time decreases as the difference between the numbers increases (Moyer & Landauer, 1967). For instance, people can judge the greater or lesser number in the pair (2, 9) more quickly than in the pair (5, 6). This well-replicated performance pattern is known as the *distance effect*. It manifests whether participants compare symbolic numbers or visuospatial numerosities like sets of dots (Buckley & Gillman, 1974). The distance effect is commonly interpreted as evidence that natural numbers are represented in part as psychophysically scaled magnitudes. It has been observed as early as infancy with numerosities (Xu & Spelke, 2000) and kindergarten with symbolic numbers (Sekuler & Mierkiewicz, 1977). Over development, children's representation of numerical magnitude changes such that response times decrease overall and the slope of the distance effect decreases (Sekuler & Mierkiewicz, 1977). These changes indicate improved precision of magnitude representations.

The magnitude representation of natural numbers and its development have also been probed using the *number line estimation* (NLE) task. In the bounded version of this task, participants are presented with a number line with the left and right endpoints labeled (e.g., as 0 and 100, respectively) and the middle segment left blank (Siegler & Opfer, 2003). Numbers are presented one at a time and in random order; the goal is to mark the position of each number on

---

these axioms often begin with 0. This reflects the subsequent development of set theory and reconstruction of the natural numbers on this foundation with 0 mapped to the empty set (e.g., Halmos, 1960; Bostock, 2009). In addition, the current ISO standard for mathematical notation defines the natural numbers as beginning with 0 (ISO, 2009). Our definition of the natural numbers follows what we take to be the modern consensus.

the number line. Performance is typically measured by average absolute error – the absolute difference between the correct position of the number and the position selected by the participant. Higher error values signify worse performance. Over development, children's accuracy on the NLE task improves gradually (Siegler, Thompson, & Opfer, 2009). There is a rich debate in the literature about the representations and processes that underlie performance on this task and its change over time. Some have proposed that children initially represent natural numbers in a compressed, logarithmically spaced fashion. Over development, this incorrect representation is thought to improve by shifting to linear spacing (Ashcraft & Moore, 2012; Opfer, Thompson, & Kim, 2016; Opfer, Siegler, & Young, 2011; Siegler & Booth, 2004; Siegler & Opfer, 2003; Thompson & Opfer, 2010). Others have argued that the representation is proportional between critical landmarks such as the midpoint (Barth & Paladino, 2011; Cohen & Blanc-Goldhammer, 2011; Slusser, Santiago, & Barth, 2013). Still others have proposed that the representation is piecewise linear between critical landmarks like place-value boundaries (Ebersbach, Luwel, Frick, Onghena, & Verschaffel, 2008; Landy, Charlesworth, & Ottmar, 2016; Landy, Silber, & Goldin, 2013; Moeller, Pixner, Kaufmann, & Nuerk, 2009). We return to these proposals in the Discussion when considering the results of the current study.

Individual differences in the magnitude representations of natural numbers, as indexed by the slope of the distance effect in the MC task, predict variation in mathematical achievement for elementary (De Smedt, Verschaffel, & Ghesquière, 2009) and middle school students (Halberda, Mazzocco, & Feigenson, 2008). Likewise, NLE performance predicts mathematical achievement in elementary school students, although not older students learning more advanced topics (Sasanguie, De Smedt, Defever, & Reynvoet, 2012). Instructional studies support a causal link between numerical magnitude representations and arithmetic skills. Young children who were tutored on the magnitudes of small natural numbers via a board game performed better on simple arithmetic problems than their peers who received a control lesson (Siegler & Ramani, 2009).

Integers

The natural numbers combined with the negative numbers form the integers {…, -2, -1, 0, 1, 2, …}. Comprehending the meaning of this number system requires engaging in additional symbolic processing. For instance, when judging the greater number in the pair (-3, -9), it is necessary to inhibit the whole number interpretations of "-3" as "3" and "-9" as "9". Children initially compare such integer pairs by using a symbolic strategy. They convert the negative numbers to natural numbers by ignoring the minus signs, comparing the numerical magnitudes, and finally reversing this judgment (Varma & Schwartz, 2011). This strategy results in a distance effect when comparing pairs of negative numbers. However, when comparing mixed integer pairs like (-4, 9), judgments can be made by noticing that positives are greater than negatives. The result is the absence of a distance effect in elementary school children.

With respect to adults' comparison of mixed integer pairs, the literature is consistent. Ganor-Stern, Tzelgov, and collaborators claimed that adults compare mixed integer pairs like children; they apply the rule that positives are greater than negatives. Their studies have largely not found distance effects under standard conditions (Ganor-Stern, Pinhas, Kallai, & Tzelgov, 2010; Tzelgov, Ganor-Stern, & Maymon-Schreiber, 2009). On the contrary, Varma and Schwartz (2011) found a developmental trend. Strategic processing of integers is eventually substituted by accessing a restructured magnitude representation that incorporates the symmetry of positives and negatives around the zero point. Adults represent the negative number line as a reflection of the positive number line, incorporating the additive inverse property $x + -x = 0$. This shift manifests as an *inverse distance effect* when comparing mixed integer pairs – as the distance between the integers increases, so does the response time. Krajcsi and Igacs (2010) detected this effect too.

As with natural numbers, patterns of estimation bias on the NLE task has been used to suggest a logarithmic-to-linear shift in the mental representation of integers. When estimating the positions of positive numbers in the range 0 – 1000 and negative numbers in the range -1000 – 0, second graders exhibit logarithmic magnitude representations. By contrast, fourth and sixth graders exhibit linear representations for both ranges (Brez, Miller, & Ramirez, 2015). This logarithmic-to-linear shift extends to larger number ranges and older people. Middle school students' estimates of integers in the negative range -10000 – 0 and the combined range -1000 – 1000 are linear, though with lower accuracy for negative numbers (Young & Booth, 2015). Adults' estimates of integers in the combined range -100 – 100 are linear and show comparable accuracy for positive and negative numbers (Ganor-Stern & Tzelgov, 2008). Performance on NLE tasks like these depends partly on recognizing the symmetry of the integers around the zero point (Saxe, Earnest, Sitabkhan, Haldar, Lewis, & Zheng, 2010).

The inverse distance effect for the MC task and symmetry effects on the NLE task led Tsang, Blair, Bofferding, and Schwartz (2015) to design a manipulative that encourages learners to incorporate the zero point in their magnitude representation of the integers. Instruction with the manipulative improved arithmetic problem-solving for elementary school students on difficult items like $3 + x = 0$.

Rational numbers

Rational numbers are the set of numbers that can be written as a ratio of two integers such that the denominator is non-zero. Magnitude comparison with symbolic fractions and decimals, as well as with non-symbolic ratios, produces the distance effect (DeWolf, Grounds, Bassok, & Holyoak, 2014; Matthews & Chesney, 2015; Matthews & Lewis, 2016; Siegler, Thompson, & Schneider, 2011; Varma & Karl, 2013). Thus, rational numbers are also thought to be integrated on the mental number line. As children begin to understand fraction magnitudes, the slope of the distance effect is nearly zero. Over time, fraction magnitude information becomes more accessible and the slope increases. By adulthood, the slope stabilizes (Gabriel, Szucs, & Content, 2013). These behavioral changes are paralleled by the discovery and use of strategies to facilitate performance. For instance, children and young adults often report using unit fractions like ½ as anchors to perform fraction magnitude comparison (Fazio, Dewolf, & Siegler, 2016; Siegler & Thompson, 2014)

In contrast to natural numbers and integers, studies using the NLE task suggest that the representation of rational numbers may be linear early on. That is, it may not undergo the logarithmic-to-linear developmental shift that has been proposed for natural numbers and integers. Both 10-year-olds and adults exhibit highly linear estimation patterns when performing NLE with decimal proportions or fractions (Iuculano & Butterworth, 2011). However, the degree of linearity does improve with development. On average, 8th graders make more accurate estimates of fractions than 6th graders; at the individual level, a higher percentage of 8th graders exhibit linear representations than 6th graders (Siegler et al, 2011).[2] These developmental gains may result in part from using unit fractions like ½ as anchors (Spinillo & Bryant, 1991). Siegler

---

[2] The failure to find a logarithmic-to-linear shift for rational numbers may be an artifact of the relatively older children who have been tested to date. It is possible that 3rd and 4th graders, who have just recently learned about fractions, will exhibit logarithmic representations. Siegler et al. (2011) argued against the expectation of logarithmic representations even in younger children.

and Thompson (2014) found that children reported using such anchors to segment the number line and estimate more accurately.

The precision of rational number magnitude representations is also associated with higher-level mathematical skills. Symbolic fraction magnitude representations predict achievement across a range of measures – fraction arithmetic, algebra, grade school standardized exams, and high school mathematical achievement (Siegler et al., 2012). Similarly, accuracy on NLE tasks using decimals predicts mathematical achievement in elementary school students (Schneider, Grabner, Zurich, & Paetsch, 2009). Finally, nonsymbolic ratio precision predicts college students' knowledge of fractions and algebra (Matthews, Lewis, & Hubbard, 2016). Causal support for the link between magnitude knowledge and arithmetic skills comes from interventions that devote more instructional time to mastering the magnitudes of fractions. These interventions show that emphasizing magnitude comprehension boosts arithmetic skills for high and low achievers alike (Fuchs et al., 2013).

Irrational numbers

Irrational numbers are incommensurable. Unlike rational numbers, they cannot be expressed as ratios of integers such that the denominator is non-zero. Their decimal expansions are infinitely long, non-repeating, and ostensibly random. When they were first proposed, the notion seemed outlandish. According to a famous anecdote, the Greek mathematician who first proved their existence was drowned at sea for challenging the ratio doctrine of numbers. It was not until the late 1800s that irrationals were formalized and properly integrated onto the real number line (Dedekind, 1963/1888). Number theorists distinguished two subclasses of irrationals – *algebraic irrationals* and *transcendental numbers*. Algebraic irrationals, like $\sqrt[2]{20}$, are the solutions to polynomial equations and can be denoted by expressions of the form $\sqrt[y]{x}$. Transcendental numbers, like π and e, are not the solutions to any polynomial equation.

Just one other study has explored whether the same human "number sense" that allows us to compare the magnitudes of natural numbers, integers, and rational numbers extends to algebraic irrationals (Obersteiner & Hofreiter, 2017). Participants performed magnitude comparisons with irrationals of the form $\sqrt[y]{x}$. Both the root *y* and the radicand *x* could vary. The researchers tested whether participants compared the magnitudes of irrationals by accessing their holistic magnitudes or by focusing on the root and radicand components. Comparisons were faster when both numbers contained a common root component, as in the pair ($\sqrt[9]{12}, \sqrt[9]{63}$). Likewise, comparisons were faster when both numbers contained a common radicand component, as in the pair ($\sqrt[10]{34}, \sqrt[13]{34}$). When both the roots and radicands differed, as in the pair ($\sqrt[10]{34}, \sqrt[15]{68}$), response times were slower.[3] Unexpectedly, response times for the common component pairs were not predicted by the distances between the root or radicand components. These mixed results require explanation. One reason might be that the irrational number stimuli were unusually complex. They do not occur in typical contexts like solving quadratic equations. Another possibility is the highly expert sample composed of mathematics graduate students and professors who may have used specialized strategies. It remains unclear how typical adults understand less complex irrational numbers like $\sqrt{2}$.

Our stimuli are also algebraic irrationals, though we focus on square root expressions of the form $\sqrt{x}$. Such expressions are often encountered when solving quadratic equations. They denote irrational numbers when the radicand *x* is not a perfect square (e.g., $\sqrt{2}$) and natural

[3] For the curious, $\sqrt[10]{34} \approx 1.422$ and $\sqrt[15]{68} \approx 1.325$. Hence, $\sqrt[10]{34} > \sqrt[15]{68}$.

numbers when the radicand is a perfect square (e.g., $\sqrt{9}$). In the following sections, we will collectively refer to both cases as *radical expressions* because both contain the radical sign.

Research questions and hypotheses

First, we asked whether radical expressions are represented as continuous magnitudes integrated on the mental number line much like natural numbers, integers, and rational numbers. We refer to this proposal as the *mental number line hypothesis*. An alternative proposal is that people use processes that capitalize on the discrete structure of radical expressions. For instance, when judging the greater or lesser number in the pair ($\sqrt{3}$, $\sqrt{8}$), people may ignore the radical signs and only compare the radicand components (3, 8) using their magnitude representations of natural numbers. This is possible because when *x* and *y* are non-negative, judgments of ($\sqrt{x}$, $\sqrt{y}$) and (*x*, *y*) are equivalent. We refer to this as the *equivalence strategy hypothesis*.

Second, we investigated whether people process irrational numbers by strategically referencing more concrete concepts like natural numbers and perfect squares. Specifically, we hypothesized the possible use of a *multiplication strategy* on the MC task whereby perfect square pairs like ($\sqrt{4}$, $\sqrt{9}$) are transformed to computationally analogous "tie" multiplication problems like 2×2 and 3×3. Such a transformation might facilitate comparison because tie problems are processed more quickly than non-tie problems of comparable size (Ashcraft, 1992; Parkman, 1972). On the NLE task, irrational numbers might be estimated in relation to perfect squares – the *landmark strategy*. For example, people may estimate the positions of irrational numbers like $\sqrt{3}$ by referencing the positions of neighboring perfect squares like $\sqrt{1}$ and $\sqrt{4}$. Finally, people may leverage their knowledge of perfect squares during arithmetic problem-solving. When simplifying $\sqrt{72}$, for instance, an inefficient strategy would be to decompose the radicand into its prime factors and then shift pairs of common factors outside the radical sign – the *prime factorization strategy*. We exemplify it by the solution procedure:

$$\begin{aligned}\sqrt{72} &= \sqrt{2\times 36}\\ &= \sqrt{2\times 2\times 18}\\ &= \sqrt{2\times 2\times 2\times 9}\\ &= \sqrt{2\times 2\times 2\times 3\times 3}\\ &= 2\sqrt{2\times 3\times 3}\\ &= 2\times 3\sqrt{2}\\ &= 6\sqrt{2}\end{aligned}$$

A more efficient approach is to factor the radicand into its largest perfect square factors and directly reduce these – the *perfect squares factorization strategy*:

$$\begin{aligned}\sqrt{72} &= \sqrt{2\times 36}\\ &= 6\sqrt{2}\end{aligned}$$

Third, we investigated whether individual differences in the mental representation and processing of irrationals explain variation in conceptual and procedural knowledge of this number system. Some have proposed that magnitude representations are the core of numerical and arithmetic performance (Link, Nuerk, & Moeller, 2014; Siegler, 2016). In contrast to this view, we hypothesize that the influence of magnitude representations on arithmetic problem-solving attenuates as the abstraction of number systems increases. Arithmetic performance with irrationals may depend less on magnitude representations and more on symbolic strategies.

Method

Participants

Overall, 81 undergraduate students from a large Midwestern university were recruited via flyers posted on campus. We posted flyers in buildings housing classrooms and in the student union. Participants' ages ranged from 18 to 24 years (M = 20.5, SD = 1.7). The sample consisted of more females than males (57 vs. 24). This is more skewed than the overall percentage of female versus male undergraduates at the university (52% vs. 48%). However, it is consistent with the percentage of females in the College of Education (61% vs. 39%), where our lab is located. On a questionnaire, 30.3% of participants self-reported that they were enrolled in a major emphasizing quantitative and analytic skills such as science, technology, engineering, mathematics, finance, or economics. This percentage is consistent with the distribution of such majors at the university. Experimental sessions spanned about 60 minutes and all participants received $12 in compensation. Our protocol was approved by the local Institutional Review Board.

Design and materials

We employed a within-subjects experimental design whereby each participant completed all levels of the four experimental tasks. See Table 1 for an overview of the design and materials. Experimental materials and data are available via the Open Science Framework (https://osf.io/6s9pc/#).

--Insert Table 1 about here --

Table 1. Overview of the design and materials used in this study.

Task	Condition	Stimuli	Number of Items	Dependent Variables
Magnitude comparison (MC)	natural numbers	0, 1, 2…, 9	204	accuracy response time
	one-digit radicals	$\sqrt{0}$, $\sqrt{1}$, $\sqrt{2}$, ..., $\sqrt{9}$	204	
Number line estimation (NLE)	natural numbers	1, 2, 3, …, 9	11	absolute error response time R^2 value
	one-digit radicals	$\sqrt{1}$, $\sqrt{2}$, $\sqrt{3}$, ..., $\sqrt{9}$	11	
	perfect squares	$\sqrt{1}$, $\sqrt{4}$, $\sqrt{9}$, ..., $\sqrt{81}$	11	
	two-digit radicals	$\sqrt{10}$, $\sqrt{20}$, $\sqrt{30}$, ... , $\sqrt{90}$	11	
Strategy questionnaire	natural numbers	How did you place 7?	2	strategy type
	one-digit radicals	How did you place $\sqrt{2}$?	2	
	perfect squares	How did you place $\sqrt{9}$?	2	
	two-digit radicals	How did you place $\sqrt{40}$?	2	
Irrationals knowledge test	number concepts	Is $\sqrt{2}$ rational or irrational?	8	accuracy strategy type
	density concepts	How many irrationals are between 0 – 1?	2	
	operation concepts	The sum of a rational and an irrational is ...?	2	
	simple arithmetic	Simplify $2\sqrt{48}$.	10	
	complex arithmetic	Simplify $\sqrt{80}+\sqrt{18}-\sqrt{20}$.	7	

Note. The stimuli for the strategy questionnaire and the irrationals knowledge test are examples from the larger set of items.

Magnitude comparison task. The first task was magnitude comparison (MC) and its stimuli were defined by four variables – type, distance, size, and perfect square status. The type variable refers to whether a number pair consisted of natural numbers or radical expressions. The natural number pairs were formed from the combinations of 0, 1, 2, …, 9 with same-number pairs excluded (90 pairs). We were particularly interested in pairs where both numbers were perfect squares, like (4, 9). Hence, we included an additional copy of each of the 12 perfect square pairs in the stimulus set. In total, 102 natural number pairs were used in each of blocks 1 and 2. Participants judged the greater of two numbers in the first block and the lesser in the second block. Natural number pairs were mirrored by 102 pairs of radical expressions formed in the same way from $\sqrt{0}$, $\sqrt{1}$, $\sqrt{2}$, …, $\sqrt{9}$. These pairs were used in blocks 3 and 4. Block 3 required greater judgments and block 4 required lesser judgments. Within each block, number pairs were presented in random order. Overall, there were 408 trials.

With respect to the distance and size variables, we used two definitions to evaluate the equivalence strategy and mental number line hypotheses (research question 1).[4] The distance

[4] The distance between the numbers being compared and the average size of the numbers being compared are indices of the precision of the magnitude representations required for successful task performance. These variables, which originate from the earliest studies of the magnitude comparison task (Moyer & Landauer, 1967; Parkman, 1971), continue to dominate studies of adults. By contrast, more recent studies of children have typically used a single variable such as the ratio of the numbers being compared (Halberda et al., 2008; Xu & Spelke, 2000). Because the current study is one of the first to investigate how *adults* understand irrational numbers, it is

variable was the absolute difference between the numbers. The first definition of distance focused on the radicand components. Thus, the distance of the pair ($\sqrt{4}$, $\sqrt{9}$) was $|4 - 9| = 5$. The second definition focused on the holistic magnitude values of the radical expressions. Thus, the distance of the pair ($\sqrt{4}$, $\sqrt{9}$) was $\left|\sqrt{4} - \sqrt{9}\right| = 1$. The size variable was defined as the average of the two numbers. The first definition again used the radicand components. Thus, the size of the pair ($\sqrt{4}$, $\sqrt{9}$) was $(4 + 9)/2 = 6.5$. The second definition again used the holistic magnitude values of the radical expressions. Thus, the size of the pair ($\sqrt{4}$, $\sqrt{9}$) was $\left(\sqrt{4} + \sqrt{9}\right)/2 = 2.5$. The perfect square variable was defined by whether both natural numbers or radicands were perfect squares like (4, 9) or ($\sqrt{4}$, $\sqrt{9}$) or not. Two dependent variables, accuracy and response time (RT) in milliseconds, were collected.

Number line estimation task. The second task was number line estimation (NLE) and it included one within-subjects variable (type) with four levels – natural numbers, one-digit radicals, perfect square radicals, and two-digit radicals. For each type, participants saw 11 stimuli. Overall, there were 44 trials. The data from the smallest stimulus and the largest stimulus were not analyzed because these mapped to one of the poles of the unmarked number line (0 and 10) in all but one case, noted below. The first block consisted of the natural numbers 0, 1, 2, …, 10. The values 0 and 10 were not analyzed. The second block consisted of the one-digit radicals $\sqrt{0}$, $\sqrt{1}$, $\sqrt{2}$, …, $\sqrt{10}$. The values $\sqrt{0}$ and $\sqrt{10}$ were not analyzed, meaning all analyzed stimuli had one-digit radicands. This is the one aforementioned case where the largest stimulus ($\sqrt{10}$) does not map to the right pole (10). The third block included the perfect squares $\sqrt{0}$, $\sqrt{1}$,

---

valuable to document whether the distance and size effects observed for other number classes extend to the irrationals. Hence, we included the distance and size variables in our analyses.

$\sqrt{4}$, …, $\sqrt{100}$. The values $\sqrt{0}$ and $\sqrt{100}$ were not analyzed. The fourth block included the two-digit radicals $\sqrt{0}$, $\sqrt{10}$, $\sqrt{20}$, …, $\sqrt{100}$. The values $\sqrt{0}$ and $\sqrt{100}$ were not analyzed, meaning all analyzed stimuli had two-digit radicands. Note that a confound exists in these stimulus sets. The values $\sqrt{1}$, $\sqrt{4}$, and $\sqrt{9}$ are shared across two types – one-digit radicals and perfect square radicals. This was unavoidable; it was forced by the distribution of perfect squares. The 11 stimuli of each type were presented in random order within a block.

Two dependent variables, average absolute error and linearity of estimates, were formed. Average absolute error was the absolute difference between a participant's selected position and the target position[5]. Linearity of estimates was the relationship between participants' selected positions and target positions as estimated by linear regression (R^2).

Strategy questionnaire. The third task was the strategy questionnaire for the NLE task. Participants completed a paper-and-pencil strategy questionnaire composed of two prompts for each of the four blocks – natural numbers, one-digit radicals, perfect square radicals, and two-digit radicals. Overall, there were 8 items. All prompts were in the form "How did you decide where to place x?", where *x* was a stimulus from one of the four blocks. See Table 1 for examples. The independent variable was number type (natural numbers, one-digit radicals,

[5] We used average absolute error (AAE) instead of percent average error (PAE). PAE is useful when participants make estimates over multiple scales (Siegler & Booth, 2004) because it allows performance to be compared between scales. Participants in the current experiment made estimates over just one scale (0-10). Therefore, we preferred the simpler AAE measure. Note that the statistical results are the same regardless of which measure is used because PAE is just a linear scaling of AAE.

perfect square radicals, and two-digit radicals) and the dependent variable was self-reported estimation strategy.

Irrationals knowledge test. The fourth and final task was the irrationals knowledge test. We constructed this novel paper-and-pencil test to measure conceptual and procedural knowledge of irrational numbers. Several studies have investigated conceptual and procedural knowledge of rational numbers and arithmetic equivalence (Alibali, Knuth, Hattikudur, McNeil, & Stephens, 2007; Booth, Newton, & Twiss-Garrity, 2014; Rittle-Johnson, Siegler, & Alibali, 2001). With these studies in mind, we operationally defined conceptual knowledge of irrational numbers as the ability to categorize particular numbers as members of this system, reason about the properties of irrationals, and relate these properties to those of other number systems. We operationally defined procedural knowledge of irrational numbers as the ability to simplify radical expressions and perform arithmetic operations on them.

Some test items were adapted from a study documenting students' misconceptions about irrational numbers (Fischbein, Jehiam, & Cohen, 1995). Conceptual items were distributed across three subsections. They required participants to classify numbers like π as either rational or irrational (8 items), answer questions about the density of rational and irrational numbers (2 items), and reason conceptually about the results of arithmetic operations on arbitrary rational and irrational numbers (2 items). Most of these items were in the selected response format. Procedural items were distributed over two subsections. Subsection 4 required performing simple arithmetic operations to reduce radical expressions (10 items) and subsection 5 required evaluating complex arithmetic expressions by performing operations on multiple radical expressions (7 items). All procedural items were fill-in-the-blank. Overall, there were 29 items. Example items from all five subsections are in Table 1. For the full test, refer to th

Supplementary Materials, section 2. Correct items earned one point each and partial credit was not awarded. Two dependent variables were constructed – accuracy on each section and arithmetic problem-solving strategies.

Procedure

After obtaining informed written consent, participants were seated alone in a quiet laboratory room in front of a Windows PC with an extended keyboard, a mouse, and a monitor measuring 55.6 cm diagonally. Participants completed the tasks in a fixed order.

Magnitude comparison task. Participants first completed the MC task, implemented in the program OpenSesame 2.9.5 (Mathôt, Schreij, & Theeuwes, 2012). In blocks 1 and 2 (natural numbers), they made greater and lesser comparisons, respectively. In blocks 3 and 4 (one-digit radicals), they made greater and lesser comparisons, respectively. Within each block, stimuli were presented in random order against a black background.

Each trial began with a blank screen for 250 ms, after which a white fixation dot was shown in the center of the screen for 1000 milliseconds. This was followed by a blank screen for 500 milliseconds. A number pair was presented next, with each number offset about 5 centimeters on either side of the screen's center. All numbers were displayed in white and in the font Cambria Math (point 60). Participants indicated the greater or lesser number by pressing the key below the target number – "Z" for the number on the left and "M" for the number on the right. The stimuli were visible until participants made a response. Feedback was then provided to discourage participants from trading accuracy for speed. Correct responses were followed by *Correct* in green and incorrect responses by *Incorrect* in red. This feedback was displayed for 500 milliseconds. Participants required about six minutes to complete each of the four blocks.

Number line estimation task and strategy questionnaire. Next, participants completed the NLE task written in JavaScript and hosted online (www.psycholab.org/FreeLab). They estimated the positions of natural numbers in block 1, one-digit radicals in block 2, perfect square radicals in block 3, and two-digit radicals in block 4. Within each block, trials were presented in random order. Participants were instructed to be as accurate as possible and that speed was irrelevant. On each trial, a natural number or radical expression was shown below a number line. These stimuli were shown in black against a white background. The number line was only labeled at the endpoints, zero on the left and ten on the right. The middle area was left blank. Participants estimated the number's position on the line by left-clicking with a computer mouse. To prevent them from rushing through the task, a blank screen was shown for 1.5 seconds between trials and clicks during this interval were ignored. To avoid learning effects, feedback was not provided. After each trial, the cursor remained where the participant last clicked. Overall, participants needed about six minutes to complete this task. Afterward, they were given the strategy questionnaire to complete at their own pace.

Irrationals knowledge test. Finally, participants completed the irrationals knowledge test without the use of electronic devices. They completed the five sections in order: (1) number concepts, (2) density concepts, (3) operation concepts, (4) simple arithmetic, and (5) complex arithmetic. To avoid learning effects, no feedback was provided. Approximately 25 minutes were allotted to respond.

Finally, demographic information was collected, participants were debriefed, and compensated with $12 in cash.

Results

Magnitude comparison task

Average accuracy across magnitude comparison blocks was high (98.2%). To determine whether there was a speed-accuracy tradeoff, we correlated the average response times and average accuracy of all participants. These variables were not significantly correlated (r = -.079, p = .486). Because of the ceiling effect in accuracy and the lack of a speed-accuracy tradeoff, subsequent analyses concentrated on the response time data by excluding all trials with incorrect responses. This meant excluding 1.8% of the overall data. Following common practice in the literature, we also excluded response times outside of the interval 200-2000 ms (Varma & Karl, 2013). This was a further 2.6% of the overall data.

Recall that the first research question asks whether the magnitudes of radical expressions are compared by referencing a mental number line or by using the equivalence strategy. The second research question, concerning the use of referential strategies, is also relevant here for answering whether perfect square radicals like ($\sqrt{4}, \sqrt{9}$) are processed more quickly via the multiplication strategy. Addressing these questions requires sensitivity to the natural confound between size and perfect square status in our stimulus set. To understand this confound, consider all natural number pairs formed by the numbers 0-9. In the range 0-9, the average is 4.5. Three of the four perfect squares in this range are less than the average (0, 1, and 4), whereas only one perfect square is greater than the average (9). As a result, the perfect square pairs were on average of smaller size than the non-perfect-square pairs. To control for this confound, an uncontrollable consequence of the non-uniform distribution of perfect squares among the natural numbers, we used a hierarchical linear regression approach. First, we estimated the effect of size. After accounting for that variance, we estimated the effect of perfect square status.

We investigated whether the canonical distance and size effects replicated for natural number comparisons and extended to radical expression comparisons. The *size effect* is the finding that after controlling for distance, response times are slower if the numbers being compared are larger (Parkman, 1971). For instance, the pair (7, 8) is compared more slowly than the pair (2, 3). Additionally, we investigated whether the type of a stimulus (natural numbers or radical expressions) and perfect square status (yes or no) affected RTs. Finally, we included the interaction between type and perfect square status in the hierarchical regression analysis. This is because if there is an advantage for comparing perfect squares, then it may only apply to radical expression pairs like ($\sqrt{1}$, $\sqrt{4}$), but not to natural number pairs like (1, 4). The dependent variable was RT and the five predictor variables were entered sequentially: distance, size, type (natural numbers or radical expressions), perfect square status (yes or no), and the type by perfect square interaction.

We addressed the first research question in two hierarchical regression analyses. The first analysis defined the distance and size variables by the radicand components of the radical expressions (e.g. ($\sqrt{1}$, $\sqrt{4}$) was converted to (1, 4)), consistent with the equivalence strategy hypothesis. The second analysis defined the distance and size variables by the radical expressions' holistic magnitude values (e.g. ($\sqrt{1}, \sqrt{4}$) was not converted), consistent with the mental number line hypothesis.

For the first analysis, the final model is shown in Table 2. Distance was entered during step 1; it was a significant predictor of RT, $R^2 = .278$, $F(1, 88) = 33.9$, $p < .001$. Size was entered during step 2 and it accounted for significant additional variance, $\Delta R^2 = .183$, $R^2 = .461$, $F(1, 87) = 29.6$, $p < .001$. Thus, we replicated the distance and size effects for natural numbers and extended them to radical expressions (Fig. 1). Number type was entered during step 3; it explained the most additional variance in RTs, $\Delta R^2 = .471$, $R^2 = .932$, $F(1, 86) = 598.5$, $p < .001$. The positive β weight means that radical expressions required more time to compare than natural numbers. Perfect square status was entered during step 4; it did not explain significant additional variance in RTs, $\Delta R^2 = .001$, $R^2 = .993$, $F(1, 85) = 1.07$, $p = .303$. The interaction between number type and perfect square value was entered during step 5; it was also a nonsignificant predictor, $\Delta R^2 = .000$, $R^2 = .993$, $F(1, 84) = .01$, $p = .935$.

--Insert Figure 1 about here --

--Insert Table 2 about here --

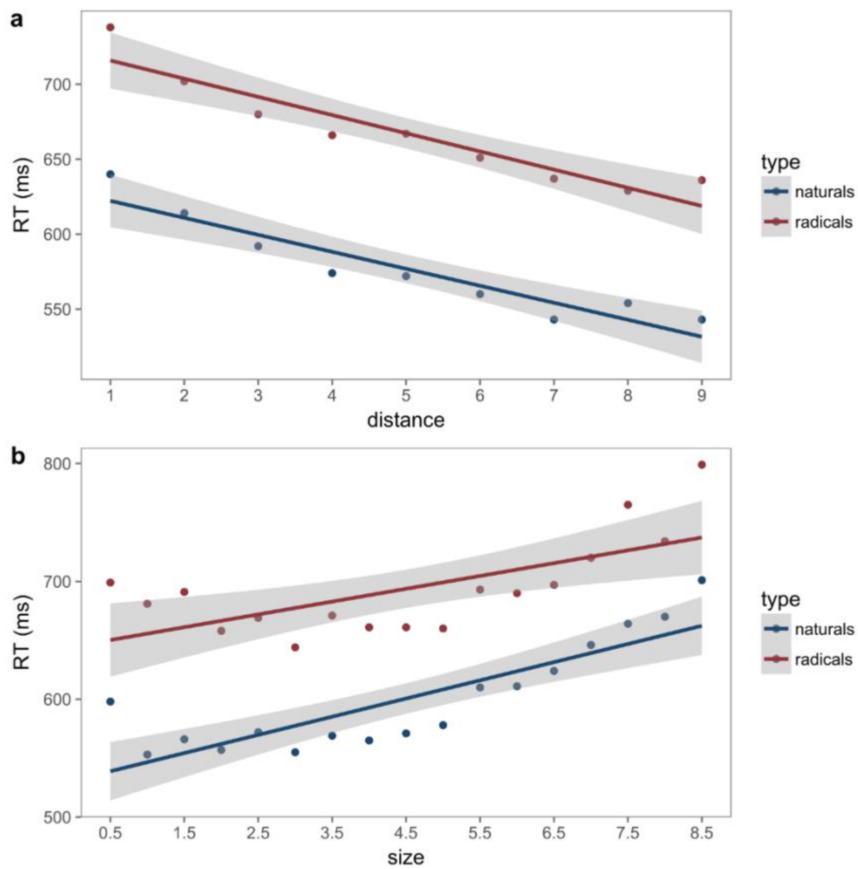

Fig. 1. Distance and size effects, defined by the radicand components and consistent with the equivalence strategy hypothesis, were found for natural number and radical expression magnitude comparisons. Shaded regions are 95% confidence envelopes.

Table 2. Final hierarchical linear regression models predicting RTs on the magnitude comparison task. The models differed on whether distance and size were defined by the radicand components (top) or the holistic magnitude values of the radical expressions (bottom).

Model	Predictor	*Std. β*	*t*	*p*	R^2
Radicand Components / Equivalence Strategy	Constant		107.49	.000	.933
	Distance	-.535	-18.40	.000	
	Size	.434	15.06	.000	
	Type	.687	22.68	.000	
	Perfect Square	-.033	-.804	.424	
	Type X Perfect Square	-.003	-.082	.935	
Holistic Magnitudes / Mental Number Line	Constant		43.03	.000	.838
	Distance	-.481	-8.99	.000	
	Size	.199	-3.74	.000	
	Type	.687	14.56	.000	
	Perfect Square	.024	.368	.714	
	Type X Perfect Square	-.003	-.053	.958	

To summarize, the first regression analysis defined the distance and size variables by the radicand components of radical expressions. The second regression analysis defined distance and size by the holistic magnitude values of the radical expressions. Distance, size, and type effects were again observed (Table 2). Critically, the first analysis explained more of the variance in RTs than the second analysis (93.3% versus 83.8%). Thus, these findings favor the equivalence strategy hypothesis over the mental number line hypothesis. Radical expressions like ($\sqrt{8}$, $\sqrt{3}$) appear to be converted to natural numbers like (8, 3) and then compared using the mental number line.

Recall that the second research question asked whether people strategically reference the more concrete natural numbers and perfect squares to process the more abstract irrationals. The superior fit of the first model (radicand components) over the second model (holistic magnitudes) provides evidence for the proposal that natural numbers are referenced to make sense of irrational numbers. However, the failure of both perfect square status and the type by perfect square status interaction to reach significance indicates that perfect squares are not referenced. That is, people do not appear to speed comparison of natural number or radical expression pairs composed of perfect squares by automatically activating associated “tie” multiplication problems. Note that there are two reasons to be cautious when interpreting this failed prediction. First, the distance, size, and type variables entered during the initial steps explained most of the variance in RTs. Little variability remained to be explained by the time the perfect square variables were entered. Second, although perfect squares may not be privileged on speeded tasks like magnitude comparison, they may benefit from special processing on unspeeded tasks like number line estimation and arithmetic problem-solving. For these reasons, we return to the second question, whether people reference perfect squares when processing irrationals, when analyzing the NLE and arithmetic problem-solving data.

To investigate individual differences, we separately estimated the first model (radicand components / equivalence strategy) for each participant. There was considerable variation between participants in the magnitudes of the β weights for the five predictors (Fig. 2), reflecting the complexity of understanding numerical processing. In later sections, we use these variables to predict individual differences in conceptual and procedural knowledge of irrational numbers.

---Insert Figure 2 about here ---

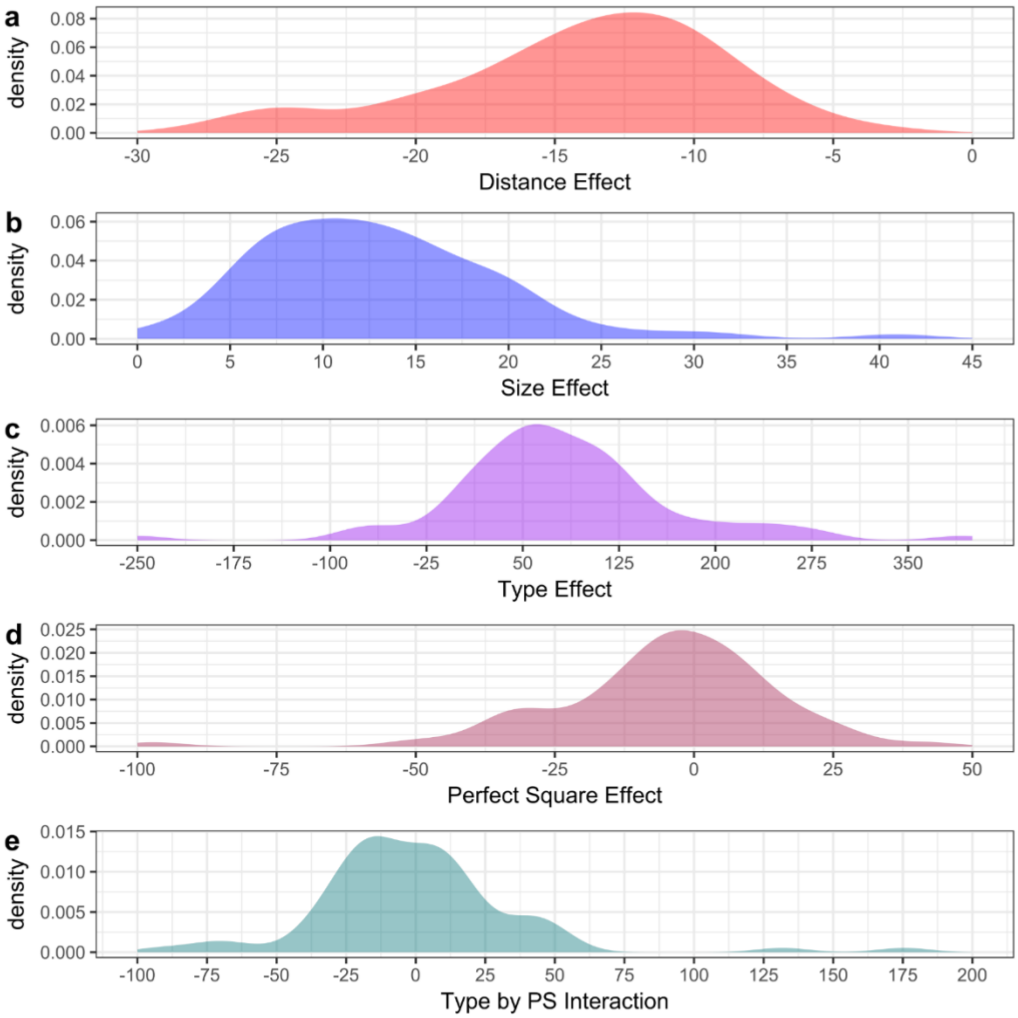

Fig. 2. Distributions of the β weights for the distance, size, type, perfect square, and type by perfect square interaction variables in linear regressions predicting each participant's magnitude comparison RTs.

Number line estimation task

Data from one participant were lost due to computer failure. Thus, the analyses included responses from 80 participants. Due to a programming error, 5.8% of the natural number trials were lost. These lost trials were random and did not affect the results. We excluded outlier responses by discarding trials for which the RT was outside of the interval 1.5 – 15 sec (4.6% of the overall data). Our analyses focused on the second research question: Do people use referencing strategies to facilitate performance?

Average absolute error and linearity. First, we analyzed average absolute error using a one-way ANOVA with type (natural numbers, one-digit radicals, perfect square radicals, and two-digit radicals) as the factor (Fig. 3). There was a main effect of type $F(3, 316) = 14.13$, $p < .001$, $\eta^2 = .118$. *Post hoc* comparisons using Tukey's procedure found that average absolute error was comparable for natural numbers and perfect squares ($p = .836$). Error for these two number types was lower than for the one-digit radicals and two-digit radicals (ps < .001), which were comparable to each other ($p = .706$).[6]

---Insert Figure 3 about here ---

[6] Although participants were instructed to focus on accuracy and not speed, a parallel analysis of response times found the same pattern of results (Supplementary Materials, section 1).

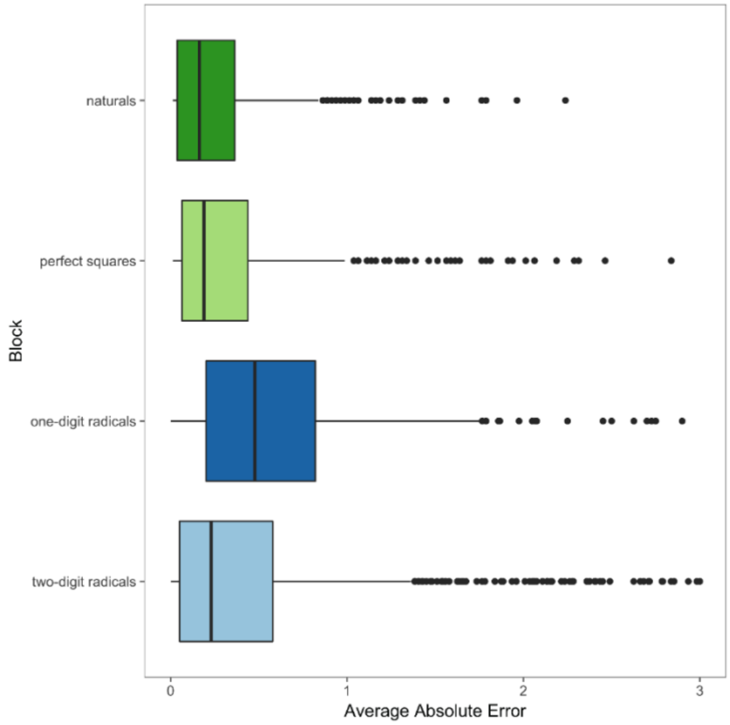

Fig. 3. Average absolute error on the number line estimation task revealed that compared to natural numbers and perfect squares, one- and two-digit radicals were more difficult to process.

How linear were participants' estimates for natural numbers, perfect squares, one-digit radicals, and two-digit radicals? For each of the four types, best-fitting lines predicting the selected position from the correct position were plotted (Fig. 4). Estimates were highly linear across all four blocks. For natural numbers and perfect squares, the R^2 values were both 1.00. For one-digit and two-digit radicals, the R^2 values were .97 and 1.00, respectively.

--Insert Figure 4 about here --

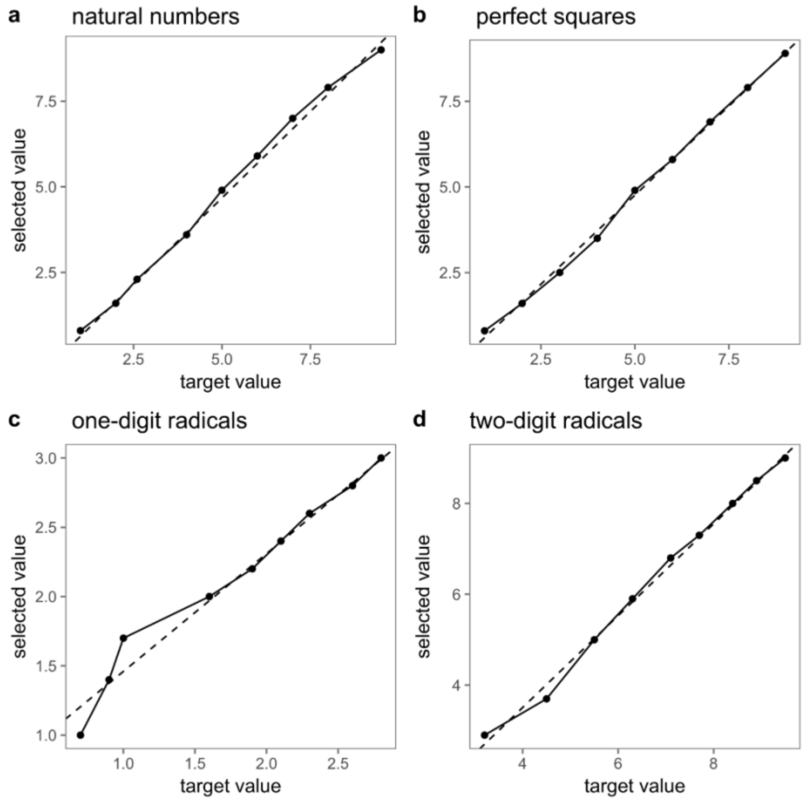

Fig.4. Number line estimation performance was highly linear on average. The dotted lines are the best-fitting regression lines.

To examine whether averaging performance concealed individual differences, we repeated them separately for each participant, estimating the best-fitting lines for each of the four number types (Fig. 5). Estimates for natural numbers and perfect squares were highly linear across participants. By contrast, there was greater variation in the linearity of participants' estimates for one-digit radicals like $\sqrt{5}$ and two-digit radicals like $\sqrt{80}$.

---Insert Figure 5 about here ---

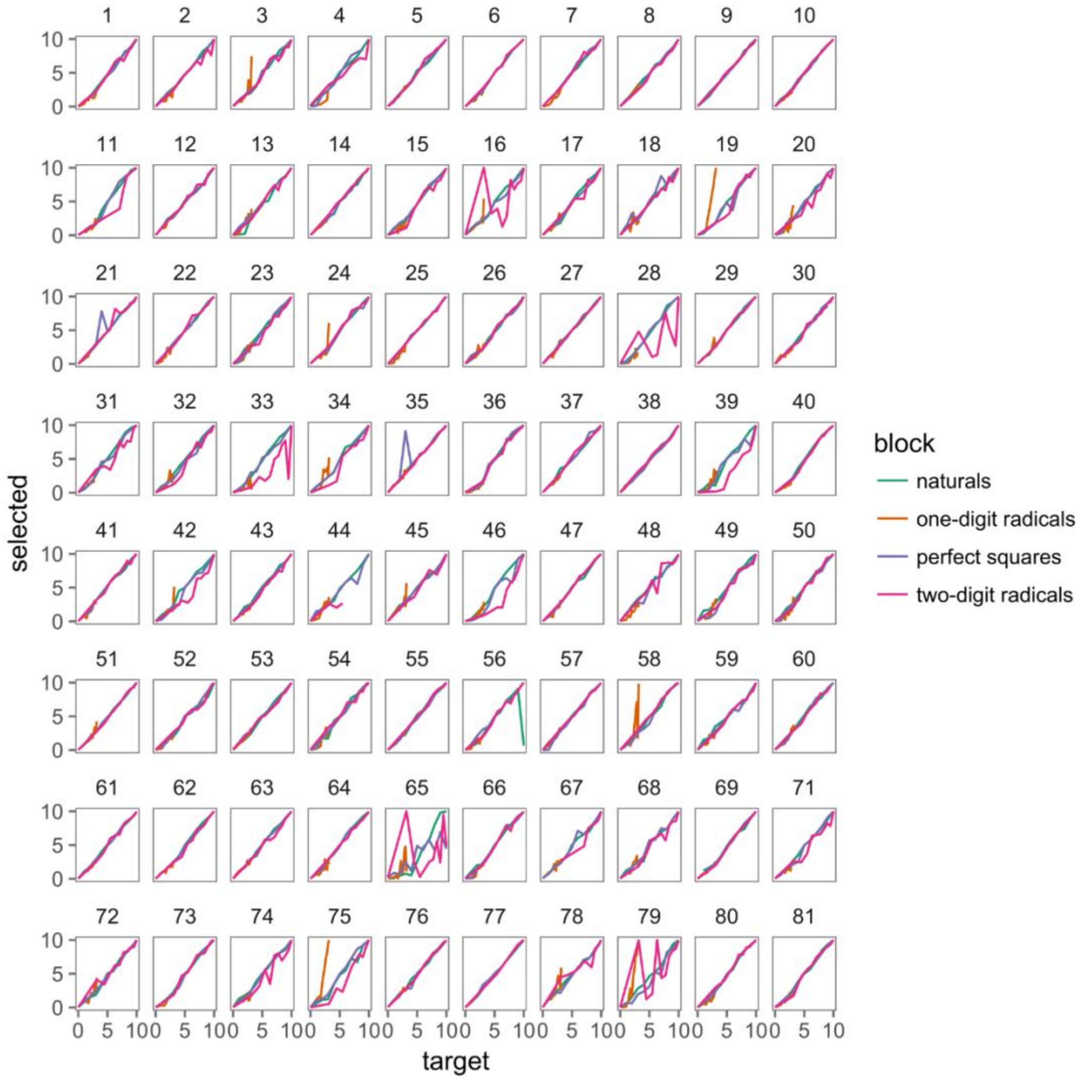

Fig. 5. Individual differences in the linearity of participants' estimates were greatest for two-digit radicals.

Estimation strategies. On the strategy questionnaire, participants self-reported the strategies they used for each block of the NLE task. After dropping illegible and incoherent responses, full data were available for 63 of 81 questionnaires. All 63 questionnaires were coded by the first author and a naïve volunteer. Interrater agreement, measured by Cohen's kappa, averaged .74 across the eight prompts. Participants reported using multiple strategies, summarized in Table 3.

--Insert Table 3 about here --

Table 3. Strategies self-reported on the number line estimation strategy questionnaire and their frequency of use.

Number Type	Strategy	Stimulus	Sample Response	% Use
Natural numbers	landmark – 0, 5, or 10	4	"It had to be before the halfway mark."	93.7
	other	7	"About ¾ of the line…"	6.3
Perfect squares	simplify & map	$\sqrt{9}$	"$\sqrt{9} = 3$ …"	95.2
	other	$\sqrt{81}$	"Guessed."	4.8
One-digit radicals	single perfect square landmark	$\sqrt{8}$	"$\sqrt{9} = 3$ near $\sqrt{9}$."	26.2
	double perfect square landmark	$\sqrt{2}$	"$\sqrt{1} = 1$, $\sqrt{4} = 2$, between $\sqrt{1}$ and $\sqrt{4}$. So between 1-2."	43.7
	other – correct	$\sqrt{2}$	"$\sqrt{2}$ is about 1.41… So it's around 1.5."	19.0
	other – incorrect	$\sqrt{8}$	"$\sqrt{8}$ is close to 8. So, estimate slightly to the left of 8."	7.9
Two-digit radicals	single perfect square landmark	$\sqrt{60}$	"I knew $\sqrt{64}$ was 8, so a little to the left."	25.4
	double perfect square landmark	$\sqrt{40}$	"I knew it was between 6 and 7 because $\sqrt{36} = 6$ and $\sqrt{49} = 7$. I put it slightly closer to where 6 was."	60.3
	other – correct	$\sqrt{40}$	"$\sqrt{40}$ is $2\sqrt{10}$. And $\sqrt{10}$ is about 3.5, so 2(3.5) is about 7."	2.4
	other – incorrect	$\sqrt{60}$	"$\sqrt{60}$ is $2\sqrt{15}$. So, estimate between 2 and 3."	8.7

For natural numbers, the most frequent strategy was to reference the landmarks 0, 5, and 10. Next, estimates were adjusted slightly to the left or right. This strategy was reported by 93.7% of participants. For example, when estimating the position of the number 3, most participants found the midpoint (5) and then shifted their estimate slightly to the left. All additional strategies were categorized as *other* (6.3%). For perfect squares like $\sqrt{9}$, most participants chose to simplify the expression to 3 and map it to the natural number line – the *simplify & map* strategy (95.2%). The use of this strategy, which is related to the equivalence strategy participants used for the MC task, suggests that perfect squares enjoy privileged processing on the NLE task. All other strategies were labeled *other* (4.8%).

For one-digit and two-digit radicals, participants continued to reference nearby landmarks during estimation. In this case, however, the landmarks were perfect squares. This finding informs the second research question concerning whether more concrete referents are used to understand the more abstract irrationals. Two main strategies were observed; we illustrate them using $\sqrt{7}$ as an example. The *single perfect square landmark* strategy was to reason that $\sqrt{7} < \sqrt{9}$ and $\sqrt{9} = 3$, so estimate slightly to the left of 3. The *double perfect square landmark* strategy was to bound the estimate from both sides. One might reason that $\sqrt{4} < \sqrt{7} < \sqrt{9}$, $\sqrt{4} = 2$, and $\sqrt{9} = 3$. Thus, one should estimate between 2 and 3. Participants were more likely to use the double landmark strategy than the single perfect square landmark strategy for two-digit radicals like $\sqrt{70}$ (60.3%) than for one-digit radicals like $\sqrt{7}$ (43.7%). Correct, but rarely used strategies were categorized as “other – correct.” Vague, incorrect, or guess responses were categorized as “other – incorrect.”

To determine whether strategy use improved accuracy when estimating the locations of one-digit radicals, a one-way ANOVA of average absolute error was conducted with strategy (single perfect square landmark, two perfect square landmarks, other – correct, other – incorrect) as a factor. We began with the first of two strategy prompts ("How did you decide where to place $\sqrt{2}$?") in the one-digit radicals section of the questionnaire. There was a main effect of strategy ($F(3, 58) = 10.75$, $p < .001$, $\eta^2 = .357$). *Post hoc* comparisons using Tukey's procedure revealed that performance was comparable for the single perfect square landmark, two perfect square landmarks, and "other – correct" strategies (ps > .05); these strategies yielded more accurate estimates than "other – correct" strategies (ps < .001). We conducted a similar one-way ANOVA for the second prompt ("How did you decide where to place $\sqrt{8}$?") and found a main effect of strategy ($F(3, 58) = 11.52$, $p < .001$, $\eta^2 = .373$). *Post hoc* comparisons again found that utilizing either of the landmark strategies or an "other – correct" strategy produced equally accurate estimates (ps > .05). These estimates were more accurate than those produced by "other – incorrect" strategies (ps < .001).

A parallel series of one-way ANOVAs was conducted for the two-digit radicals prompts. For the first prompt ("How did you decide where to place $\sqrt{40}$?"), a main effect of strategy was detected ($F(3, 58) = 20.15$, $p < .001$, $\eta^2 = .510$). *Post hoc* comparisons revealed that average absolute error was comparable when using a single perfect square landmark, two perfect square landmarks, and "other – correct" strategies (ps > .05). Error was worse when using "other – incorrect" strategies (ps < .001). For the second prompt ("How did you decide where to place $\sqrt{60}$?"), a main effect of strategy was detected again $F(3, 58) = 6.75$, $p < .001$, $\eta^2 = .259$) and *post hoc* comparisons revealed that the perfect square and "other – correct" strategies were comparable to each other. Both of these strategies were superior to "other – incorrect" strategies (ps < .01).

To summarize, participants who used the single or double perfect square landmark strategies were no more accurate than participants who used “other – correct” strategies. This was not because participants who used these strategies were at floor, as their estimates were more accurate than those of participants who used “other – incorrect” strategies.

Irrationals knowledge test

All 81 participants completed a paper-and-pencil irrationals knowledge test composed of two sections – conceptual knowledge and procedural knowledge. These were in turn composed of five subsections – number concepts, density concepts, operation concepts, simple arithmetic, and complex arithmetic. Recall that Table 1 lists example items from each subsection. The full test is in the supplementary materials (section 2). Table 4 summarizes performance across all sections and subsections.

---Insert Table 4 about here ---

Table 4. Performance on all sections and subsections of the irrationals knowledge test.

Section	Subsection	Mean (%)	SD
Conceptual	Number concepts	75.1	20.4
	Density concepts	49.0	26.4
	Operation concepts	66.9	22.6
average		*63.6*	*17.2*
Procedural	Simple arithmetic	61.0	27.0
	Complex arithmetic	35.8	35.2
average		*48.4*	*29.3*

To determine the number of constructs being measured across these subsections, we conducted a principal components analysis using varimax rotation. The two procedural subsections loaded onto a single *procedural* factor (eigenvalue = 2.3, 46.0% of variance explained) and the three conceptual subsections loaded onto a single *conceptual* factor (eigenvalue = 1.1, 22.8% of variance explained).[7] Hence, subsequent analyses focused on average scores on the conceptual section and the procedural section.

To address the second research question, which concerns the use of referencing strategies during arithmetic problem-solving, we coded the strategies used on the arithmetic section of the irrationals knowledge test. Interrater agreement, measured by Cohen's kappa, averaged .95. We focused on whether participants used inefficient or efficient strategies.

Some participants used the inefficient *prime factorization strategy* exclusively (17%). It involves solving arithmetic problems by decomposing radicands into prime factors and then shifting pairs of common factors outside the radical sign. See Table 5 for an example. By contrast, 40% of participants used the *perfect squares factorization strategy* at least once. It involves directly identifying the largest perfect square factors of the radicand. This chunking strategy enabled participants to solve problems using fewer operations (Table 5).

---Insert Table 5 about here ---

[7] See Figure S2 in the Supplementary Materials for a plot of the two factors.

Table 5. Example of applying the two strategies to solve the complex arithmetic problem $\sqrt{48}+\sqrt{75}-\sqrt{27}$ from the irrationals knowledge test. Using the prime factorization strategy requires computing and maintaining more partial products than using the perfect squares factorization strategy.

Strategy	Step	Partial Products	Use
Prime Factorization	0	$\sqrt{48}+\sqrt{75}-\sqrt{27}$	17%
	1	$\sqrt{48}=\sqrt{2\times 24}$	
	2	$\sqrt{48}=\sqrt{2\times 2\times 2\times 2\times 3}$	
	3	$\sqrt{48}=2\sqrt{2\times 2\times 3}$	
	4	$\sqrt{48}=4\sqrt{3}$	
	5	$\sqrt{75}=\sqrt{3\times 25}$	
	6	$\sqrt{75}=\sqrt{3\times 5\times 5}$	
	7	$\sqrt{75}=5\sqrt{3}$	
	8	$\sqrt{27}=\sqrt{3\times 9}$	
	9	$\sqrt{27}=\sqrt{3\times 3\times 3}$	
	10	$\sqrt{27}=3\sqrt{3}$	
	11	$4\sqrt{3}+5\sqrt{3}-3\sqrt{3}=6\sqrt{3}$	
	12		
Perfect Squares Factorization	0	$\sqrt{48}+\sqrt{75}-\sqrt{27}$	40%
	1	$\sqrt{48}=\sqrt{16\times 3}$	
	2	$\sqrt{48}=4\sqrt{3}$	
	3	$\sqrt{75}=\sqrt{25\times 3}$	
	4	$\sqrt{75}=5\sqrt{3}$	
	5	$\sqrt{27}=\sqrt{9\times 3}$	
	6	$\sqrt{27}=3\sqrt{3}$	
	7	$4\sqrt{3}+5\sqrt{3}-3\sqrt{3}=6\sqrt{3}$	

To evaluate whether arithmetic strategy use affected participants' procedural scores on the irrationals knowledge test, a one-way ANOVA was conducted. The strategy factor had four levels: prime factorization, perfect squares factorization, both strategies, and other. The category "other" refers to participants who used incorrect strategies or guessed. There was a main effect of strategy on procedural scores ($F(3, 77) = 29.71$, $p < .001$, $\eta^2 = .537$). *Post hoc* comparisons with Tukey's procedure revealed that procedural scores were comparable between participants who used the prime factorization strategy and those who used the perfect squares factorization strategy ($p = .260$). The only significant score differences were between participants who used either of these strategies (or both) and participants who used "other" strategies (*p*s < .001). This indicates that participants who used the less efficient prime factorization strategy were not at floor, as they outperformed those who used "other" strategies (incorrect procedures). The perfect squares strategy is not associated with higher accuracy on complex arithmetic problems than the use of the prime factorization strategy. However, it may confer other advantages such as reducing processing time and working memory demand. We could not test these possibilities because we did not measure these dependent variables.

Predicting individual differences in knowledge of irrationals

The third research question concerned whether individual differences in the magnitude representation and strategic processing of radical expressions predicts conceptual and procedural knowledge of irrationals. To answer this question, we extracted two sets of predictor variables from each participant, as described above. The first set consisted of the β weights for the three significant variables (distance, size, and type) when predicting response times on the magnitude comparison task. The β weights for the distance and size variables index the precision of magnitude representations and are common predictors of mathematical achievement in children

(De Smedt et al., 2009). The β weight for the type variable indicates the additional time required to compare radical expressions versus natural numbers. In other words, it reflects the time needed to apply the equivalence strategy. The second set of predictor variables were the average absolute errors when estimating the positions of one-digit radicals and two-digit radicals on the NLE task. These variables reflect the efficiency of referencing neighboring perfect squares. The magnitude comparison variables were entered first and the NLE variables were entered second in separate hierarchical linear regressions predicting accuracy on the conceptual section and the procedural section.

For the conceptual knowledge measure, the magnitude comparison variables added during the first step were nonsignificant (R^2 = .074, F(3, 76) = 2.02, p = .119). Adding the NLE variables during the second step accounted for significant additional variance (ΔR^2 = .115, R^2 = .188, F(2, 74) = 5.23, p = .007). In the final model, shown in Table 6, the magnitude comparison variable type and the NLE variable two-digit radicals were significant predictors.

---Insert Table 6 about here ---

Table 6. Final models predicting conceptual and procedural knowledge of irrationals.

Measure	Predictor	*Std. β*	*t*	*p*	R^2
Conceptual Knowledge	Constant		11.87	.000	.188
	MC – β Distance	.039	.271	.788	
	MC – β Size	.053	.370	.712	
	MC – β Type	.235	2.21	.030	
	NLE 1D Radicals	-.022	-.184	.855	
	NLE 2D Radicals	-.355	-2.97	.004	
Procedural Knowledge	Constant		8.28	.000	.322
	MC – β Distance	.125	.945	.348	
	MC – β Size	-.045	-.343	.732	
	MC – β Type	.151	1.55	.125	
	NLE 1D Radicals	-.104	-.952	.344	
	NLE 2D Radicals	-.411	-3.76	.000	

For the procedural knowledge measure, the magnitude comparison variables accounted for significant variance in step 1 (R^2 = .140, $F(3, 76)$ = 4.14, p = .009), and the NLE variables for significant additional variance in step 2 (ΔR^2 = .182, R^2 = .322, $F(2, 74)$ = 9.91, p < .001). In the final model, only the NLE variable two-digit radicals was significant (Table 6).

These findings inform the second and third research questions. With respect to the second research question, participants' ability to estimate the spatial locations of two-digit radicals (all irrational numbers) predicted conceptual and procedural scores. This indicates that strategically referencing perfect squares is associated with success on higher-level tests of mathematical knowledge. Concerning the third research question, neither the distance nor the size variables from the magnitude comparison task predicted knowledge of irrationals in the final models. This suggests that these variables, which index the precision of magnitude representations, are less important when predicting performance on unspeeded tasks measuring higher-level mathematical knowledge. Further, the significance of the magnitude comparison variable type for predicting conceptual scores supports the role of the equivalence strategy. Understanding when judgments of irrational numbers can and cannot be simplified to judgments of natural numbers may depend on conceptual knowledge of number systems.

Discussion

Summary

To study how people understand highly abstract mathematical concepts, we used irrational numbers as a test case. Specifically, we addressed three research questions that tested multiple hypotheses. The first question asked whether radical expressions are compared by referencing a mental number line or by using the equivalence strategy. We addressed this question by fitting two models to the magnitude comparison RT data. The first model, consistent with the equivalence strategy, defined the distance and size variables by the radicands' magnitudes. The second model, consistent with the mental number line strategy, defined the distance and size variables by the holistic magnitudes of the radical expressions. The equivalence strategy model fit the data best. In addition, the type variable, which indicates whether the stimuli being compared are natural numbers or radical expressions, was a significant predictor above and beyond the distance and size variables. Taken together, these results suggest that people compare the magnitudes of radical expressions by referencing their natural number components (i.e., radicands).

The second research question asked whether people reference perfect squares to make sense of irrational numbers. Examples of such referencing, besides the equivalence strategy, are evident on the NLE task. Strategy self-reports suggest that people reference nearby perfect squares to better estimate the locations of one-digit and two-digit irrationals – the single and double landmark strategies. Referencing strategies were also observed on the complex arithmetic subsection of the irrationals knowledge test. When simplifying radical expressions, 40% of participants factored radicands into perfect squares. This may be a more efficient strategy than the prime factorization strategy used by 17% of participants.

The third research question asked whether individual differences in the speed and accuracy with which irrational numbers are processed on the simple MC and NLE tasks predicts individual differences in higher-level conceptual and procedural knowledge measures comparable to classroom mathematics tests. The type variable for the magnitude comparison task, which indicates whether the stimuli are natural numbers or radical expressions, was a significant predictor of conceptual knowledge. We interpret this as reflecting the importance of selectively accessing the natural number components of radical expressions (i.e., radicands and roots) when reasoning about irrationals. Accuracy with two-digit radicals on the NLE task was also a significant predictor. It explained both conceptual and procedural knowledge. This may reflect the cognitive value of referencing perfect squares when processing irrational numbers.

Although magnitude representations are necessary for arithmetic competence, they are not as critical as for natural numbers, integers, and rational numbers. The role of magnitude representations appears to attenuate in importance as the abstraction of number systems increases. Hence, alternative mechanisms must be invoked to explain expertise with irrationals.

Theoretical implications – a referential processing account

Some theories of numerical and arithmetic development emphasize the recruitment, repurposing, and refinement of perceptuo-motor and visuospatial systems (Blair, Tsang, & Schwartz, 2013; Landy, Allen, & Zednik, 2014; Siegler, 2016). Based on our findings that natural numbers and perfect squares support an understanding of the irrationals on a range of tasks, we hypothesize that abstract mathematical concepts are understood by selectively recruiting more concrete representations that align with task demands and stimulus properties. We call these representations *referents* and the recruitment and alignment processes *referencing* or *referential processing*.

Mathematical tasks like magnitude comparison, number line estimation, and arithmetic problem-solving can impose high computational demands on limited cognitive resources. Referential processing reduces these demands in task- and stimulus-specific ways. Magnitude comparison with radical expressions, for instance, can be performed using a “guess and check” algorithm to determine the greater or lesser value in number pairs like $(\sqrt{7}, \sqrt{4})$. Due to the substantial computational demands of using this algorithm, people choose to use the more efficient equivalence strategy instead. Radical expressions are interpreted as natural number referents like (7, 4) because the judgments are the same in both cases. On the NLE task, the computational demands of estimating the positions of one-digit and two-digit radicals are also high. They are circumvented by locating estimates with respect to neighboring perfect square referents – the single and double perfect square landmark strategies. Likewise, referential processing supports arithmetic problem-solving on the irrationals knowledge test by enabling expressions like $\sqrt{48}$ to be simplified efficiently via perfect squares factorization:

$$\sqrt{48} = \sqrt{16 * 3} = 4\sqrt{3}.$$

This is less computationally demanding as it requires fewer steps and partial products than the prime factorization strategy:

$$\sqrt{48} = \sqrt{2 * 24} = \sqrt{2 * 2 * 12} = \sqrt{2 * 2 * 2 * 6} = \sqrt{2 * 2 * 2 * 2 * 3} = 2\sqrt{2 * 2 * 3}$$
$$= 4\sqrt{3}$$

Critically, the referential processing account states that referents are task- and stimulus-specific. On the magnitude comparison task, for instance, the appropriate referents were natural numbers. In contrast, the appropriate referents on the NLE and arithmetic problem-solving tasks were perfect squares. To explain why particular concrete mathematical concepts are appropriate on some tasks but not others, it is useful to focus on the alignment between what a task demands and what a referent provides. An appropriate referent simplifies a task without altering its underlying structure. For example, the equivalence strategy aligns well with the magnitude comparison task because the comparisons ($\sqrt{x}$, $\sqrt{y}$) and (x, y) have the same response when x and y are non-negative. In this regard, the equivalence strategy's focus on the natural number components of radical expressions constitutes a virtuous form of the *whole number bias* (Ni &

Zhou, 2005).[8] Referential processing, then, can reduce computational demands and improve performance while preserving task structure.

The extent to which referential processing enhances performance relative to default strategies depends on the task to be performed. On the MC task, referential processing improves how quickly irrational numbers are processed. During NLE, it improves accuracy compared to “other-incorrect” strategies. It is important to acknowledge that referential processing does not always improve accuracy. On the arithmetic problem-solving test, for example, the use of perfect square referents was *not* associated with higher accuracy compared to prime number referents. The benefits afforded in this context appear to be more modest and may be restricted to the speed domain, a prediction that the current study did not test.

Referential processing across number systems and mathematical tasks

Our referential processing account unifies a constellation of seemingly disparate phenomena in the mathematical cognition literature. The prominence of referents has been demonstrated using more open-ended prompts, in naturalistic studies, and in formal analyses of number naming systems. In production studies, where participants must generate numbers in a specified range like 1 – 100, the frequency with which numbers are produced follows a power law. Smaller numbers are produced more frequently than larger numbers. Critically, the exceptions are referents in the base-10 system like 10 and 100 or the larger culture like 12, 15, and 50. These numbers are produced more frequently than the power law predicts (Noma & Baird, 1975). Corpora studies reveal the same pattern; referents in the base-10 system like 10 and

---

[8] We thank an anonymous reviewer for suggesting this connection between referential processing and the whole number bias.

100 are produced more frequently than expected given the otherwise power-law distribution of numbers (Dehaene & Mehler, 1992). In a corpus study and formal analysis of number words in the Dutch language, Pollman and Jansen (1996) found that referents in the base-10 system like the base raised to an integer power (e.g., $10^1 = 10$) are among the "favourite numbers." This finding was extended to Swedish by Ericsson, Biley, and Geary (2010).

Studies using the MC and NLE tasks also support the use of referential processing across a range of number systems (Table 7). For natural numbers, referential processing has been observed when comparing one-digit numbers, with end-anchoring to zero (Banks, 1977; Pinhas & Tzelgov, 2012). Comparisons of multi-digit numbers are made with reference to intervening multiples of 10 (Brysbaert, 1995; Franklin, Jonides, & Smith, 2009). The question of whether referential processing is used when estimating the position of a natural number on a bounded number line (and if so, under which conditions) is a contentious one. We focus here on findings and models consistent with referential processing.[9] Ebersbach at al. (2008) proposed that the largest number familiar to children serves as a referent during NLE. Moeller et al. (2009) proposed that before children have integrated the ones and tens places in their understanding of number, 10 serves as a referent when estimating numbers in the range 0 – 100. Landy et al. (2013) proposed that scale words like 'million' serve as referents when estimating across very large ranges like 1 thousand – 1 billion. Across these studies, variants of a bilinear model with a discontinuity at the proposed referent (the largest familiar number, 10, and 1 million,

[9] Evidence exists against the referential processing account and in favor of continuous performance (whether logarithmic or linear) on the NLE task (Ashcraft & Moore, 2012; Opfer et al., 2011; Siegler & Booth, 2004; Siegler & Opfer, 2003; Opfer et al. 2016).

respectively) fit the data better than a linear or a logarithmic model. Finally, some have applied cyclical models to account for the NLE performance of children and adults (Barth & Paladino, 2011; Cohen & Blanc-Goldhammer, 2011; Slusser et al., 2013). For our purposes, these models can be characterized as positing referents at endpoints, midpoints, and quartiles. They also posit proportional estimation between referents in opposition to the linear estimation of the models just reviewed.

Even natural number arithmetic can be supported by base-10 referents. For instance, one-digit addition problems such as 9 + 8 can be solved by decomposition to the referent 10: 9 + 8 = 9 + (1 + 7) = (9 + 1) + 7 = 10 + 7 = 17 (Baroody, Purpura, Eiland, Reid, & Paliwal, 2015). For multi-digit arithmetic, the referents are often multiples of 10, as when 49 + 38 is transformed to 50 + 37 (Torbeyns, de Smedt, Ghesquiere, & Verschaffel, 2009).

---Insert Table 7 about here ---

Table 7. Across number systems and tasks, people use many kinds of referents flexibly.

	Naturals		Integers	Rationals		Irrationals
	single-digit	multi-digit		fractions	decimals	
Magnitude Comparison	zero	multiples of 10^n	zero	unit fractions, decimals	naturals	naturals[a]
Number Line Estimation	endpoints midpoints quartiles	multiples of 10^n, midpoints, quartiles	zero	unit fractions, decimals	quartiles tenths	perfect squares[b]
Arithmetic Computation	10	multiples of 10^n	zero	?	?	perfect squares[c]

[a] Supported by the equivalence strategy on the MC task.

[b] Supported by the single and double perfect square landmark strategies on the NLE task.

[c] Supported by the perfect squares factorization strategy on the irrationals knowledge test.

Referential processing also extends to more abstract number systems. With integers, zero can be used as a referent when performing magnitude comparison, number line estimation, number bisection, and arithmetic computation (Gullick & Wolford, 2013; Tsang et al., 2015; Tsang & Schwartz, 2009; Varma & Schwartz, 2011). Fractions are often processed by referencing unit fractions like ½ when making magnitude comparison and number line estimation judgments (Fazio et al., 2016; Siegler & Thompson, 2014). When rational numbers are expressed as decimal proportions, their magnitudes can be referenced with respect to natural numbers (Varma & Karl, 2013). Finally, during number line estimation with decimals, quartiles like 0.25 and tenths like 0.10 are often referenced (Houseworth, 2016).

Instructional implications

Our results have three potential implications for instructional design in mathematics. First, performance patterns on the MC and NLE tasks suggest that our sampled undergraduates represent the magnitudes of irrationals accurately. Nonetheless, performance on the NLE task reveals that some people place irrationals, denoted by one-digit expressions (e.g., $\sqrt{2}$) and two-digit expressions (e.g., $\sqrt{40}$), on the number line inaccurately. It might be possible to improve their performance by giving students explicit strategy instruction. They could be guided to estimate the values of irrationals by referencing perfect squares via the single and double perfect square landmark strategies.

The second pedagogical implication concerns strategy differences during arithmetic problem-solving. Participants who used the prime factorization strategy required more steps and likely spent more time on task. Those who used the perfect squares factorization strategy, in contrast, could skip steps by directly accessing the perfect square factors under the radical signs. Not all learners discover such strategies on their own, suggesting that perfect squares factorization should be taught during formal instruction. Doing so may improve students' adaptive expertise, or their ability to recognize and apply optimal strategies (Hatano & Oura, 2003).

Third, it is worth noting that the correlation between magnitude knowledge and arithmetic expertise is weaker for irrationals than other number systems. Interventions that improve the magnitude representations of children result in greater arithmetic expertise for natural numbers, integers, and rational numbers (Fuchs et al., 2013; Siegler & Ramani, 2009; Tsang et al., 2015). Our individual differences analyses, however, did not find a relationship between the precision of magnitude representations as measured by the distance and size effects and conceptual and procedural knowledge of irrationals. This suggests that training magnitude representations may not produce similar learning gains for the irrationals. We tentatively suggest that a better approach for teaching more abstract mathematical concepts like the irrationals is to focus on strategies for relating them to referents according to stimulus properties and task demands.

Limitations

This study was constrained in several ways. First, the NLE task involved collecting strategy prompts after the entire task ended rather than after each trial in a block. This is potentially problematic because strategies can change on a trial-by-trial basis, as has been shown for fractions (Fazio et al., 2016). Future studies should collect strategies after every trial. Relatedly, only two strategy prompts were presented for each block. It is possible that participants used different strategies for the unprompted items and that these strategies were not captured in the current study. These limitations could be addressed in a future study that focuses on the NLE task for irrational numbers and collects strategy prompts after every trial.

A second limitation is that much of the variation in conceptual and procedural knowledge of irrational numbers was unexplained. Unexplained variance in the conceptual knowledge measure might reflect differential understanding of the cardinalities of infinite sets such as the rationals, irrationals, and reals. This knowledge is crucial for defining and reasoning about these number systems. Unexplained variance in the procedural knowledge measure might reflect differential fluency in the multiplication and division skills necessary for factoring and simplifying radical expressions. Future studies should investigate these and other possible explanations.

Conclusion

Most numerical cognition research focuses on relatively concrete concepts like natural numbers, integers, and rational numbers. Our study of irrationals elucidates how the mind grasps abstract mathematical concepts incapable of being directly perceived or fully grounded in visuospatial referents. We suggest that the numerical magnitudes and symbolic procedures associated with irrationals are understood by referencing natural numbers and perfect squares. Such referential processing may explain performance across a wide range of tasks and number systems (Table 7), supporting an intuitive understanding of otherwise alien concepts. Thus far, we have only learned that experts use referential processing to understand irrationals. Our next step is to explore the development of referential processing in novices and how it supports the discovery, comprehension, and instruction of diverse mathematical content. Doing so will illuminate how the abstract becomes concrete.

References

Alibali, M. W., Knuth, E. J., Hattikudur, S., McNeil, N. M., & Stephens, A. C. (2007). A longitudinal examination of middle school students' understanding of the equal sign and equivalent equations. *Mathematical Thinking and Learning*, *9*, 221-247. https://doi.org/10.1080/10986060701360902

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition*, *44*(1), 75–106. https://doi.org/10.1016/0010-0277(92)90051-I

Banks, W. P. (1977). Encoding and processing of symbolic information in comparative judgments. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 11, pp. 101–159). New York, NY: Academic Press.

Baroody, A. J., Purpura, D. J., Eiland, M. D., Reid, E. E., & Paliwal, V. (2015). Does fostering reasoning strategies for relatively difficult basic combinations promote transfer by K-3 students? *Journal of Educational Psychology*, *108*(4), 576–591. https://doi.org/10.1037/edu0000067

Bender, A., Schlimm, D., & Beller, S. (2015). The Cognitive Advantages of Counting Specifically: A Representational Analysis of Verbal Numeration Systems in Oceanic Languages. *Topics in Cognitive Science*, *7*(4), 552–569. https://doi.org/10.1111/tops.12165

Blair, K., Tsang, J., & Schwartz, D. (2013). The bundling hypothesis - How perception and culture give rise to abstract mathematical concepts in individuals. In S. Vosniadou (Ed.), *International Handbook of Research on Conceptual Change* (2nd ed., pp. 322–340). New York City: Routledge.

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development*, *79*(4), 1016–1031. https://doi.org/10.1111/j.1467-

8624.2008.01173.x

Brez, C. C., Miller, A. D., & Ramirez, E. M. (2015). Numerical estimation in children for both positive and negative numbers. *Journal of Cognition and Development*, *17*(2), 341–358. https://doi.org/10.1080/15248372.2015.1033525

Brysbaert, M. (1995). Arabic number reading: On the nature of the numerical scale and the origin of phonological recoding. *Journal of Experimental Psychology: General*, *124*(4), 434–452. https://doi.org/10.1037/0096-3445.124.4.434

Buckley, P. B., & Gillman, C. B. (1974). Comparisons of digits and dot patterns. *Journal of Experimental Psychology*, *103*(6), 1131–1136. https://doi.org/10.1037/h0037361

Cohen, D., & Blanc-Goldhammer, D. (2012). Numerical bias in bounded and unbounded number line tasks. *Psychonomic Bulletin & Review*, *18*(2), 331–338. https://doi.org/10.3758/s13423-011-0059-z.

De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology*, *103*(4), 469–479. https://doi.org/10.1016/j.jecp.2009.01.010

Dedekind, J. (1963). *Essays on the Theory of Numbers*. *WW Beman, Open Court, Chicago* (1st ed.). New York City: Dover Publications, Inc.

DeWolf, M., Grounds, M. A., Bassok, M., & Holyoak, K. J. (2014). Magnitude comparison with different types of rational numbers. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1), 71–82. https://doi.org/10.1037/a0032916

Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology*, *123*(1), 53–72. https://doi.org/10.1016/j.jecp.2014.01.013

Fazio, L. K., Dewolf, M., & Siegler, R. S. (2016). Strategy use and strategy choice in fraction magnitude comparison. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, *42*(1), 1–16. https://doi.org/10.1037/xlm0000153

Fischbein, E., Jehiam, R., & Cohen, D. (1995). The Concept of Irrational Numbers in High-School Students and Prospective Teachers. *Educational Studies in Mathematics*, *29*(1), 29–44. Retrieved from http://www.jstor.org/stable/3482830

Franklin, M. S., Jonides, J., & Smith, E. E. (2009). Processing of order information for numbers and months. *Memory & Cognition*, *37*(5), 644–654. https://doi.org/10.3758/MC.37.5.644

Fuchs, L. S., Schumacher, R. F., Long, J., Namkung, J., Hamlett, C. L., Cirino, P. T., … Changas, P. (2013). Improving at-risk learners' understanding of fractions. *Journal of Educational Psychology*, *105*(3), 683–700. https://doi.org/10.1037/a0032446

Gabriel, F. C., Szucs, D., & Content, A. (2013). The development of the mental representations of the magnitude of fractions. *PLoS ONE*, *8*(11), 1–14. https://doi.org/10.1371/journal.pone.0080016

Ganor-Stern, D., Pinhas, M., Kallai, A., & Tzelgov, J. (2010). Holistic representation of negative numbers is formed when needed for the task. Q*uarterly Journal of Experimental Psychology*, *63*, 1969–1981. https://dx.doi.org/10.1080/17470211003721667

Ganor-Stern, D., & Tzelgov, J. (2008). Negative numbers are generated in the mind. *Experimental Psychology*, *55*, 157–163. https://dx.doi.org/10.1027/1618-3169.55.3.157

Gullick, M. M., & Wolford, G. (2013). Understanding less than nothing: Children's neural response to negative numbers shifts across age and accuracy. *Frontiers in Psychology*, *4*(548), 1–17. https://doi.org/10.3389/fpsyg.2013.00584

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

Hatano, G., & Oura, Y. (2003). Commentary : Reconceptualizing school learning using insight from expertise research. *Educational Researcher*, *32*(8), 26–29. https://doi.org/10.3102/0013189X032008026

Houseworth, J. (2016). *The role of rational numbers in mathematical achievement and decision making*. University of Minnesota - Twin Cities. Retrieved from http://login.ezproxy.lib.umn.edu/login?url=http://search.proquest.com.ezp3.lib.umn.edu/docview/1870036427?accountid=14586

Huber, S., Klein, E., Willmes, K., Nuerk, H.-C., & Moeller, K. (2014). Decimal fraction representations are not distinct from natural number representations - evidence from a combined eye-tracking and computational modeling approach. *Frontiers in Human Neuroscience*, *8*(April), 172. https://doi.org/10.3389/fnhum.2014.00172

Iuculano, T., & Butterworth, B. (2011). Understanding the real value of fractions and decimals. *Quarterly Journal of Experimental Psychology (2006)*, *64*(11), 2088–98. https://doi.org/10.1080/17470218.2011.604785

Krajcsi, A., & Igács, J. (2010). Processing negative numbers by transforming negatives to positive range and by sign shortcut. *European Journal of Cognitive Psychology*, *22*, 1021–1038. https://dx.doi.org/10.1080/09541440903211113

Link, T., Nuerk, H.-C., & Moeller, K. (2014). On the relation between the mental number line and arithmetic competencies. *Quarterly Journal of Experimental Psychology (2006)*, *67*(8), 1597–613. https://doi.org/10.1080/17470218.2014.892517

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: an open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–24. https://doi.org/10.3758/s13428-011-0168-7

Matthews, P. G., & Chesney, D. L. (2015). Fractions as percepts? Exploring cross-format distance effects for fractional magnitudes. *Cognitive Psychology*, *78*, 28–56. https://doi.org/10.1016/j.cogpsych.2015.01.006

Matthews, P. G., & Lewis, M. R. (2016). Fractions we cannot ignore: The nonsymbolic ratio congruity effect. *Cognitive Science*, 1–19. https://doi.org/10.1111/cogs.12419

Matthews, P. G., Lewis, M. R., & Hubbard, E. M. (2016). Individual differences in nonsymbolic ratio processing predict symbolic math performance. *Psychological Science*, *27*(2), 191–202. https://doi.org/10.1177/0956797615617799

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*(5109), 1519–1520. https://doi.org/10.1038/2151519a0

Obersteiner, A., & Hofreiter, V. (2017). Do we have a sense for irrational numbers? *Journal of Numerical Cognition*, *2*(3), 170–189. https://doi.org/10.5964/jnc.v2i3.43

Opfer, J. E., Thompson, C. A., & Kim, D. (2016). Free versus anchored numerical estimation: A unified approach. *Cognition*, *149*, 11–17. https://doi.org/10.1016/j.cognition.2015.11.015

Parkman, J. M. (1972). Temporal aspects of simple multiplication and comparison. *Journal of Experimental Psychology*, *95*(2), 437–444. https://doi.org/10.1037/h0033662

Pinhas, M., & Tzelgov, J. (2012). Expanding on the mental number line: Zero is perceived as the “smallest”. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1187–1205. https://dx.doi.org/10.1037/a0027390

Sasanguie, D., De Smedt, B., Defever, E., & Reynvoet, B. (2012). Association between basic numerical abilities and mathematics achievement. *British Journal of Developmental Psychology*, *30*(2), 344–357. https://doi.org/10.1111/j.2044-835X.2011.02048.x

Saxe, G. B., Earnest, D., Sitabkhan, Y., Haldar, L. C., Lewis, K. E., & Zheng, Y. (2010). Supporting generative thinking about the integer number line in elementary mathematics. *Cognition and Instruction*, *28*(4), 433–474. https://doi.org/10.1080/07370008.2010.511569

Schneider, M., Grabner, R. H., Zurich, E., & Paetsch, J. (2009). Mental number line, number line estimation, and mathematical achievement: Their interrelations in grades 5 and 6. *Journal of Educational Psychology*, *101*(2), 359–372. https://doi.org/10.1037/a0013840

Sekuler, R., & Mierkiewicz, D. (1977). Children’s judgments of numerical inequality. *Child Development*, *48*(2), 630–633. https://dx.doi.org/10.2307/1128664

Siegler, R. S. (2016). Magnitude knowledge: The common core of numerical development. *Developmental Science*, *19*(3), 341–361. https://doi.org/10.1111/desc.12395

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., … Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, *23*(7), 691–697. https://doi.org/10.1177/0956797612440101

Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games—but not circular ones—improves low-income preschoolers’ numerical understanding. *Journal of Educational Psychology*, *101*(3), 545–560. https://doi.org/10.1037/a0014239

Siegler, R. S., & Thompson, C. A. (2014). Numerical landmarks are useful - except when they’re not. *Journal of Experimental Child Psychology*, *120*(1), 39–58. https://doi.org/10.1016/j.jecp.2013.11.014

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology*, *62*(4), 273–296. https://doi.org/10.1016/j.cogpsych.2011.03.001

Spinillo, A. G., & Bryant, P. (1991). Children's proportional judgments : The importance of "Half." *Child Development*, *62*(3), 427–440. https://dx.doi.org/10.2307/1131121

Thompson, C. A., & Opfer, J. E. (2010). How 15 hundred is like 15 cherries: Effect of progressive alignment on representational change in numerical cognition. *Child Development*, *81*, 1768–1786. https://doi.org/10.1111/j.1467-8624.2010.01509.x

Torbeyns, J., de Smedt, B., Ghesquiere, P., & Verschaffel, L. (2009). Jump or compensate? Strategy flexibility in the number domain up to 100. *ZDM - International Journal on Mathematics Education*, *41*(5), 581–590. https://doi.org/10.1007/s11858-009-0187-3

Tsang, J. M., Blair, K. P., Bofferding, L., & Schwartz, D. L. (2015). Learning to "see" less than nothing: Putting perceptual skills to work for learning numerical structure. *Cognition and Instruction*, *33*(2), 154–197. https://doi.org/10.1080/07370008.2015.1038539

Tsang, J. M., & Schwartz, D. L. (2009). Symmetry in the semantic representation of integers. In N. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st annual conference of the cognitive science society* (pp. 323–328). Austin, TX: Cognitive Science Society.

Tzelgov, J., Ganor-Stern, D., Maymon-Schreiber, K. (2009). The representation of negative numbers: Exploring the effects of mode of processing and notation. *The Quarterly Journal of Experimental Psychology*, *62*, 605-624. https://dx.doi.org/10.1080/17470210802034751

Varma, S., & Karl, S. R. (2013). Understanding decimal proportions: Discrete representations, parallel access, and privileged processing of zero. *Cognitive Psychology*, *66*(3), 283–301. https://doi.org/10.1016/j.cogpsych.2013.01.002

Varma, S., & Schwartz, D. L. (2011). The mental representation of integers: An abstract-to-concrete shift in the understanding of mathematical concepts. *Cognition*, *121*(3), 363–385. https://doi.org/10.1016/j.cognition.2011.08.005

Verguts, T., & Van Opstal, F. (2005). Dissociation of the distance effect and size effect in one-digit numbers. *Psychonomic Bulletin & Review*, *12*(5), 925–930. https://doi.org/10.3758/BF03196787

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), 1–11. https://doi.org/10.1016/S0010-0277(99)00066-9

Young, L. K., & Booth, J. L. (2015). Student magnitude knowledge of negative numbers. *Journal of Numerical Cognition*, *1*(1), 38–55. https://doi.org/10.5964/jnc.v1i1.7

Zhang, J., & Norman, D. A. (1995). A representational analysis of numeration systems. *Cognition*, *57*(3), 271–295. https://doi.org/10.1016/0010-0277(95)00674-3

Figure captions

Fig. 1. Distance and size effects, defined by the radicand components and consistent with the equivalence strategy hypothesis, were found for natural number and radical expression magnitude comparisons. Shaded regions are 95% confidence envelopes.

Fig. 2. Distributions of the β weights for the distance, size, type, perfect square, and type by perfect square interaction variables in linear regressions predicting each participant's magnitude comparison RTs.

Fig. 3. Average absolute error on the number line estimation task revealed that compared to natural numbers and perfect squares, one- and two-digit radicals were more difficult to process.

Fig. 4. Number line estimation performance was highly linear on average. The dotted lines are the best-fitting regression lines.

Fig. 5. Individual differences in the linearity of participants' estimates were greatest for two-digit radicals.