
Volume 121, Issue 3, December 2011
ISSN 0010-0277
pp. 289–466
ELSEVIER
COGNITION

Cognition 121 (2011) 363–385

Contents lists available at SciVerse ScienceDirect

Cognition

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

The mental representation of integers: An abstract-to-concrete shift in the understanding of mathematical concepts

Sashank Varma [a,*], Daniel L. Schwartz [b]

[a] University of Minnesota, Department of Educational Psychology, 165 Education Sciences Building, 56 East River Rd., Minneapolis, MN 55455, USA
[b] Stanford University, School of Education, Stanford, CA 94305-3096, USA

ARTICLE INFO

Article history:
Received 2 April 2010
Revised 2 August 2011
Accepted 8 August 2011
Available online 21 September 2011

Keywords:
Integers
Analog magnitudes
Mental number line
Symbolic distance effect
Inverse distance effect
Size effect
Semantic congruence effect
SNARC effect
Formal models

ABSTRACT

Mathematics has a level of structure that transcends untutored intuition. What is the cognitive representation of abstract mathematical concepts that makes them meaningful? We consider this question in the context of the integers, which extend the natural numbers with zero and negative numbers. Participants made greater and lesser judgments of pairs of integers. Experiment 1 demonstrated an *inverse distance effect*: When comparing numbers across the zero boundary, people are faster when the numbers are near together (e.g., −1 vs. 2) than when they are far apart (e.g., −1 vs. 7). This result conflicts with a straightforward symbolic or analog magnitude representation of integers. We therefore propose an *analog-x* hypothesis: Mastering a new symbol system *restructures* the existing magnitude representation to encode its unique properties. We instantiate analog-*x* in a *reflection model:* The mental negative number line is a reflection of the positive number line. Experiment 2 replicated the inverse distance effect and corroborated the model. Experiment 3 confirmed a developmental prediction: Children, who have yet to restructure their magnitude representation to include negative magnitudes, use rules to compare negative numbers. Taken together, the experiments suggest an abstract-to-concrete shift: Symbolic manipulation can transform an existing magnitude representation so that it incorporates additional perceptual-motor structure, in this case symmetry about a boundary. We conclude with a second *symbolic-magnitude model* that instantiates analog-*x* using a feature-based representation, and that begins to explain the restructuring process.

1. Introduction

Mathematics has a level of structure that transcends untutored intuition. Cantor's theory of transfinite numbers enables us to speak of infinity, a "number" we never encounter, and even of a hierarchy of infinities. Mathematics also describes empirical phenomena that cannot be directly experienced. For example, Einstein's theory of general relativity uses the counterintuitive geometry of Riemann to characterize the curvature of space–time. What are the cognitive underpinnings of abstract mathematical concepts that make them meaningful?

The integers, which extend the natural numbers with zero and negative numbers, are abstract numbers.[1] They are meaningful for adults despite the fact that they do not have ready perceptual-motor referents. One does not handle negative physical objects, and zero is arguably the prototype of abstraction, a form without substance. It was only a few 100 years ago that mathematicians properly formalized integers. For these reasons, the mathematician Felix Klein (1925) claimed that with integers, "for the first time, we

* Corresponding author. Tel.: +1 612 625 6718; fax: +1 612 624 8241.
E-mail addresses: sashank@umn.edu (S. Varma), daniel.schwartz@stanford.edu (D.L. Schwartz).

0010-0277/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2011.08.005

[1] Mathematicians variously define the natural numbers as the positive integers $\{1,2,3,\ldots\}$ or the non-negative integers $\{0,1,2,3,\ldots\}$. We use the term to denote one or the other class depending on the local context.

meet the transition from concrete to formal mathematics. The complete mastery of this transition requires a higher-order ability in abstraction" (p. 23). The meaningfulness of integers for adults makes them a useful test case for investigating how people come to understand abstract mathematical concepts that are unlikely to be innate or induced from everyday experiences.

Integer understanding is also relevant for the cognitive science debate between symbolic and embodied representations in higher-order cognition (Anderson, 1978; Kosslyn, 1981; Landy & Goldstone, 2007; Miller & Johnson-Laird, 1976; Pylyshyn, 1981). Symbolic theories characterize the acquisition of mathematical concepts as the compilation of symbolic rules into efficient and automatic procedures (Anderson, 1982; Newell, 1990). Affiliated instructional methods emphasize the mastery of symbolic rules (Anderson, Corbett, Koedinger, & Pelletier, 1995; Kaminski, Sloutsky, & Heckler, 2008). By contrast, embodied theories describe mathematics learning as abstraction over perceptual-motor experience (Barsalou, 1999; Glenberg, 1997; Johnson, 1987; Lakoff & Núñez, 2000; Varela, Thompson, & Rosch, 1992). Associated instructional methods often involve grounding symbolic notations, structures, and rules in hands-on activities (Montessori, 1966).

Here, we propose a different learning progression – an abstract-to-concrete shift. Symbolic rules *restructure* perceptual-motor representations to serve higher order cognition. When learning integers, children already have an analog magnitude representation of natural number that exhibits perceptual-motor properties (Moyer & Landauer, 1967; Sekuler & Mierkiewicz, 1977). Because negative numbers and zero do not have a ready perceptual-motor basis, children initially understand them by using symbolic rules that map them to natural numbers. With experience, the structure of the symbolic rules transforms the original magnitude representation of natural number to directly embody the unique properties of the integers, such as the fact that zero is a boundary between positives and negatives. By this restructuring hypothesis, symbolic rules provide a mechanism for articulating perceptual-motor representations into richer and more abstract concepts.

1.1. From the natural numbers to the integers

Natural numbers can be mentally represented as analog magnitudes. A key piece of evidence is that the time it takes to judge which of two natural numbers is greater exhibits a continuous logarithmic function that is also found when comparing physical quantities, such as the loudness of two tones. In a seminal study, Moyer and Landauer (1967) had participants judge which of two natural numbers was greater. Participants exhibited a *symbolic distance effect*: They were faster comparing numbers that were far apart (1 vs. 9) than near together (1 vs. 4). The distance effect is commonly interpreted as evidence for a mental number line such that magnitudes that are farther apart on the line are easier to discriminate (Restle, 1970). A related finding is the *size effect*: For pairs of natural numbers a fixed distance apart, people are faster comparing smaller numbers (1 vs. 4) than larger numbers (6 vs. 9) (Parkman, 1971). The size effect indicates that the number line is psychophysically scaled, because smaller magnitudes are easier to discriminate than larger magnitudes.

Two hypotheses have been advanced in the literature for how the natural numbers are extended to the integers, which we refer to as *symbol*+ and *analog*+. Symbol+ proposes that the magnitude representation of natural numbers is supplemented with symbolic rules for handling zero and negative numbers. (The '+' denotes symbols plus magnitudes.) This has been referred to as the 'phylogenetic' hypothesis (Fischer, 2003), the 'magnitude-polarity' hypothesis (Shaki & Petrusic, 2005), and the 'components' representation (Ganor-Stern & Tzelgov, 2008; Tzelgov, Ganor-Stern, & Maymon-Schreiber, 2009). The resulting hybrid representation is shown in Table 1. By this account, when comparing positive numbers (1 vs. 9), people consult their natural number line. The farther apart the numbers, the faster the judgment. When comparing negative numbers, people use symbolic rules to strip signs and invert the predicate (e.g., which of −1 vs. −9 is greater becomes which of 1 vs. 9 is lesser). They then use the natural number line to compare the magnitudes (e.g., that $1 < 9$ implies that $-1 > -9$). When comparing a positive number and a negative number (−1 vs. 7), people apply rules such as "positives are greater than negatives," and there is no need to consult magnitude representations. Therefore, they should not exhibit distance effects for mixed comparisons of a positive number and a negative number.

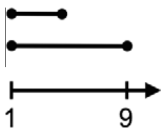
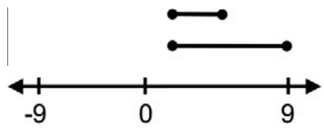
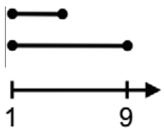
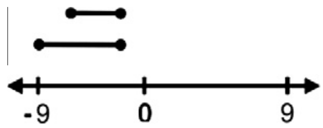
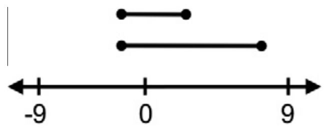
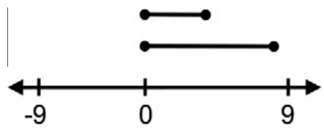
Analog+ proposes that the magnitude representation of natural numbers is extended by adding magnitude representations of zero and negative numbers to the mental number line. This hypothesis has been called the 'ontogenetic' hypothesis (Fischer, 2003), the 'number-line' hypothesis (Shaki & Petrusic, 2005), and the 'holistic' representation (Ganor-Stern & Tzelgov, 2008; Tzelgov et al., 2009). A spatial model of the analog+ hypothesis is shown in Table 1. By this account, the mental number line has the same topological organization as the conventional number line of mathematics. Thus, unlike symbol+, analog+ predicts that all comparisons of integers should show distance effects, including mixed comparisons.

This paper hypothesizes a third representation of integers we call *analog-x*. The '*x*' denotes a transformation of the magnitude representation of natural numbers. Critically, this representation includes additional structure reflecting the unique properties of the integer symbol system relative to the natural number symbol system, such as the *additive inverse* property: For every integer x, there exists an integer $-x$ such that $x + (-x) = 0$. The human capacity to transform analog magnitude representations to incorporate symbolic structure is what makes it possible to have semantic, perceptual-motor representations of abstract mathematical concepts. We present one instantiation of analog-*x* after the results of Experiments 1 are in hand, and another in Section 6.

1.2. Prior studies of integers

The literature reports two broad classes of behavioral effects for integer comparisons – magnitude effects and

Table 1
Symbol+ and analog+ distance effect predictions for different comparison types.

Comparison	Distance		Symbol+		Analog+	
	Far	Near	Process	Product	Process	Product
Positive	1 vs. 9	1 vs. 4	Judge magnitudes		Judge magnitudes	
Negative	−9 vs. −1	−4 vs. −1	(1) Strip signs (2) Invert predicate[a] (3) Judge magnitudes	−9 vs. −1 → 9 vs. 1 −4 vs. −1 → 4 vs. 1 > → < 	Judge magnitudes	
Mixed	−1 vs. 7	−1 vs. 2	Positives > negatives	7 > −1 2 > −1	Judge magnitudes	
Zero	0 vs. 8	0 vs. 3	Positives > zero Zero > negatives	8 > 0 3 > 0	Judge magnitudes	

[a] Example assumes a greater comparison.

judgment effects. The effects are defined with respect to the different types of integer comparisons (Table 1). *Positive* comparisons involve two natural numbers. *Negative* comparisons involve two negative numbers. *Mixed* comparisons involve a positive and a negative. If the positive has the greater absolute value, then the comparison is *mixed-positive*, and if the negative is greater, then it is *mixed-negative*. Finally, there are two types of *zero* comparisons: zero vs. a positive (*zero-positive*), and zero vs. a negative (*zero-negative*).

1.2.1. Magnitude effects

A magnitude effect is one where the magnitudes (i.e., absolute values) of numbers influence comparison time, *per* the distance and size effects. They are of primary importance because they are the basis for differentiating symbol+ and analog+. Both hypotheses predict magnitude effects for positive comparisons because both propose that people directly consult a positive number line. Additionally, both hypotheses predict magnitude effects for negative comparisons, but for different reasons. Analog+ proposes that people directly consult the negative number line. Symbol+ proposes that people indirectly consult a number line representation, first using rules to handle the negative signs and then the natural number line to judge their absolute magnitudes. Because a mental number line is consulted in both cases, the finding of a distance effect for negative comparisons (Tzelgov et al., 2009; Krajcsi & Igács, 2010) does not differentiate the hypotheses. And, because mental number lines are psychophysically scaled, the finding of a size effect for negative comparisons does not differentiate the hypotheses (Ganor-Stern & Tzelgov, 2008; Shaki & Petrusic, 2005).

By contrast, mixed and zero comparisons distinguish the hypotheses. Symbol+ proposes these comparisons are completely handled by rules; for example, "positives are greater than negatives." Thus, there should be no effect of magnitude for these comparison types. By contrast, analog+ proposes that mixed and zero comparisons are handled with an extended number line, so there should be an effect of magnitude. Prior studies of these comparisons types have produced conflicting results. Tzelgov et al. (2009) did not find a distance effect for mixed comparisons, consistent with symbol+. By contrast, Krajcsi and Igács (2010) found an *inverse distance effect* for mixed comparisons. People were faster for near mixed comparisons (e.g., −1 vs. 5) than far mixed comparisons (e.g. −9 vs. 5). This finding is inconsistent with both symbol+ and analog+. Finally, Fischer and Rottman (2005) found a distance effect for zero-positive comparisons, but not for zero-negative comparisons. The distance effect for zero-positive comparisons is consistent with both analog+ and symbol+ if one assumes that zero is part of the natural number line and hence treated as a magnitude. However, the absence of a distance effect for zero-negative comparisons is consistent only with symbol+.

1.2.2. Judgment effects

Judgment effects occur when number magnitudes are associated with cognitive or visuo-motor processes. The *semantic congruence effect* depends on the cognitive process of judging greater or lesser. When comparing large natural numbers (6 vs. 9), people make greater judgments faster than lesser judgments. When comparing small natural numbers (1 vs. 4), people make lesser judgments faster than greater judgments (Banks, Fujii, & Kayra-Stuart, 1976;

Holyoak, 1978). This is taken as evidence that the right and left poles of the mental number line for natural numbers are associated with the greater and lesser predicates, respectively.

Symbol+ predicts a semantic congruence effect for positive comparisons. It does not predict a semantic congruence effect for negative comparisons because it proposes that people do not represent negative magnitudes directly. Consistent with symbol+, Ganor-Stern and Tzelgov (2008) and Tzelgov et al. (2009) found a semantic congruence effect for positive comparisons, but not for negative comparisons.

Analog+ differs from symbol+ under the assumption that positive numbers are associated with greater judgments and negative numbers with lesser judgments. There should be a semantic congruence effect for positive comparisons – greater judgments should be faster than lesser judgments – and this should extend to all comparisons that are positive dominant (mixed-positive and zero-positive, where the positive number has the greater absolute value). And, there should be an *inverse* semantic congruence effect for negative comparisons – lesser judgments should be faster than greater judgments – and this should extend to all comparisons that are negative dominant (mixed-negative and zero-negative, where the negative number has the greater absolute value). Consistent with analog+, Shaki and Petrusic (2005) found a semantic congruence effect for positive comparisons and an inverse semantic congruence effect for negative comparisons.

A second judgment effect depends on the association between number magnitude and the left and right sides of space. The Spatial-Numerical Association of Response Codes (SNARC) effect is the finding that people respond faster on the left side of space when judging small natural numbers and faster on the right side when judging large natural numbers (Dehaene, Bossini, & Giraux, 1993; Fischer, 2003).[2] This suggests that the mental number line has the conventional left-to-right orientation for positive numbers. Using a parity judgment task – deciding odd or even – Fischer and Rottman (2005) found an *inverse* SNARC effect for negative numbers. People made left responses faster when judging large negative numbers (e.g., -1,) and made right responses faster when judging small negative numbers (e.g., -9). This is consistent with symbol+, which proposes that people understand negative numbers by stripping their signs and mapping them to positive numbers (e.g., the large negative number -1 is mapped to the small positive number 1). By contrast, Shaki and Petrusic (2005) found a standard SNARC effect, with positive comparisons faster with right responses and negative comparisons faster with left responses. This is consistent with analog+ under the assumption that positive magnitudes are associated with right responses and negative magnitudes with left responses.

1.2.3. Explaining the inconsistencies

Both the magnitude and judgment data exhibit inconsistencies across experiments – sometimes supporting symbol+, sometimes analog+, and sometimes neither hypothesis. Rather than discounting the results or the hypotheses, it is possible that people switch between representations depending on contextual or task factors. For example, even if people have a magnitude representation of negative numbers, as proposed by analog+, it is easy to imagine that given a steady stream of mixed comparisons, they would eventually ignore the number magnitudes and follow the rule "positives are greater than negatives," as proposed by symbol+. This might explain why Tzelgov et al. (2009) found no effect of distance for mixed comparisons. That study was designed to evaluate the relative contributions of intentional vs. automatic processing, and therefore included thousands of trials, more than half of which were mixed comparisons. The preponderance of mixed trials may have encouraged people to apply the rule of looking at the signs of numbers and ignore their magnitudes, and thus behave in concordance with symbol+.

Contextual factors may also be responsible for conflicting judgment effects. For example, Ganor-Stern and Tzelgov (2008) used two-digit numbers instead of one-digit numbers to study the semantic congruence effect. Nuerk, Weger, and Willmes (2001) demonstrated that two-digit numbers are not represented as simple magnitudes, but rather are broken down into tens and ones components through symbolic operations. This symbolic processing may have primed participants to adopt a symbol+ representation. There are also contextual differences between studies finding conflicting SNARC effects. Fischer and Rottman (2005) used a parity judgment task, whereas Shaki and Petrusic (2005) used a number comparison task. In a clear demonstration of the effect of context, Shaki and Petrusic (2005) showed that the direction of the SNARC effect depends on whether positive and negative comparisons are blocked or intermixed.

If contextual factors influence the number representations that people employ for a particular task, then the search should not be for the single representation that underlies all performance, but rather for the range of representations that people have at their disposal. Our proposal is that adults have an analog-*x* representation of integers that they can consult, but that children, who have had much less experience with negative numbers, have not developed an analog-*x* representation, and therefore reason using a symbol+ representation.

1.3. Plan for this paper

The purpose of the current research is not to choose between symbol+ and analog+, but rather to investigate the hypothesis that adults also have an analog-*x* representation that encodes more mathematical structure than is available in natural numbers. We focus on the magnitude effects, particularly the effect of distance for mixed comparisons, as it is most diagnostic. We demonstrate that relatively

[2] Although the conventional interpretation of the SNARC effect is that it reflects the direct association between number magnitudes and the left and right sides of space, recent research suggests that this association is indirect, mediated by the categories +polarity and -polarity (Proctor & Cho, 2006) or the verbal labels "large" and "small" (Gevers et al., 2010; Santens & Gevers, 2008). Regardless of whether the association is direct or indirect, the SNARC effect remains a judgment effect because it is not a function of number magnitudes themselves, as is the case for the distance and size effects.

experienced adults can consult an analog-*x* representation, whereas relatively inexperienced children cannot, even though they can make integer comparisons effectively.

Experiment 1 demonstrates an inverse distance effect for mixed comparisons (Krajcsi & Igács, 2010; Varma, Schwartz, Lindgren, & Go, 2007). Adults are faster at mixed comparisons when the numbers are near together (−1 vs. 2) than when they are far apart (−1 and 7). This is inconsistent with both symbol+ and analog+, and leads to the attribution of the additional structure of analog-*x*. We instantiate analog-*x* in a formal model that proposes that the negative number line is a reflection of the positive number line. This model accounts for the inverse distance effect for mixed comparisons, as well as the other findings of Experiment 1. Experiment 2 uses a different participant population and different experimental paradigm to ensure the generalizability of the model's fit.

Experiment 3 evaluates the developmental claim that children have not yet restructured their number representation into analog-*x*, and therefore use the symbol+ representation. We conclude by offering a second formal model that instantiates analog-*x* using a feature-based representation. This model begins to explain the process by which symbolic rules transform analog representations to incorporate additional structure. It represents an important step in reconciling theories that emphasize either continuous or discrete aspects of number, but not both.

2. Experiment 1

Experiment 1 investigated the adult mental representation of integers. Highly educated adults viewed two integers and judged which was greater (or lesser). All four integer comparison types were intermixed within each block to minimize the adoption of task-dependent strategies that would apply to one comparison type but not others (e.g., "always choose the positive number"). For each comparison type, judgments were made at two distances, near and far, which provides the relevant contrast for evaluating the distance effect. Mixed comparisons (positives vs. negatives) are of particular interest because symbol+ predicts that there should be no effect of distance – near and far judgments should be equally fast. By contrast, analog+ predicts that there should be a standard distance effect, with near judgments slower than far judgments. Experiment 1 also evaluated the size, semantic congruence, and SNARC effects.

2.1. Method

2.1.1. Participants

Twenty-one graduate students (16 female; 2 left-handed) from the Stanford University School of Education participated for course credit. All reported normal or corrected-to-normal vision.

2.1.2. Stimuli and design

The integers were grouped into five sets: small natural numbers {1,2,3,4}, large natural numbers {6,7,8,9}, small negative integers {−1,−2,−3,−4}, large negative integers {−6,−7,−8,−9}, and zero {0}. We define the size of an integer x by its absolute value – small if $|x| < 5$ and large if $|x| > 5$. By this definition, −1 and 2 are small integers, and −8 and 9 are large. This simplifies the analysis of the size effect below.

The stimuli were pairs of integers that formed three fully crossed within-subjects factors: comparison type, distance, and predicate. The first factor, comparison type, had four levels: positive, negative, mixed, and zero. There were also sub-factors (mixed-positive, mixed-negative, zero-positive, and zero-negative) that we describe in the Results section.

The second factor, comparison distance, had two levels: near and far. Distance was defined as the mathematical distance of the integers being compared ($|x - y|$). not as the absolute distance ($||x| - |y||$). This distinction is only relevant for mixed comparisons, for example the mathematical distance between −1 vs. 2 is 3, whereas the absolute distance is 1. Near-distance stimuli had a distance of 2 or 3, and far-distance stimuli had a distance of 7 or 8. For positive, negative, and zero comparisons, near-distance stimuli were formed by pairing two small integers (including zero) or pairing two large integers. Far-distance stimuli were formed by pairing a small integer (including zero) with a large integer. Mixed comparisons were formed differently. Near-distance stimuli paired a small positive integer and a small negative integer (−1 vs. 2). Far-distance stimuli paired a small negative integer and a large positive integer (−1 vs. 7) or vice versa (−7 vs. 1).

This composition scheme equated the mathematical distance of the near and far levels across the four comparison types. It also ensured that distance and size were not confounded for mixed comparisons. Specifically, if far-distance stimuli were formed by pairing large integers (−8 vs. 9), and if near-distance stimuli were formed by pairing small integers (−1 vs. 2), then it would be impossible to determine whether response time differences between near and far stimuli represent a distance effect or a size effect.

The five sets of integers and foregoing constraints yielded 44 stimuli. Each appeared twice within each experimental block, once in each left–right order (1 vs. 9, 9 vs. 1), for a total of 88 stimuli per block. Stimuli were randomly ordered within each block. There were four experimental blocks.

The third factor, comparison predicate, had two levels: choose the greater or lesser number. It was varied across blocks to minimize participant confusion within a block. Participants were randomly assigned to one of two predicate orders (> < > < and < >< >).

2.1.3. Procedure

The stimuli appeared on a PC running E-Prime with a CRT monitor measuring 40.25 cm diagonally. Using a 5-button response box, participants placed a left-hand finger on the leftmost button and a right-hand finger on the rightmost button. Participants completed practice blocks of 12 greater and 12 lesser judgments sampled from each cell of the design. They then completed the four experimental blocks. Each block began with 12 unmarked practice trials to allow participants to switch into making greater or lesser judgments.

On each trial, "Ready" appeared in 18-point Courier font in the center of the screen for 1000 ms, followed by a fixation cross for 1000 ms, followed by the pair of numbers, which appeared two spaces to the left and right of center. Positive numbers were presented without a "+" sign. Participants pressed the button below the greater (or lesser) number. After a response or a 3000 ms deadline, the stimulus was replaced with feedback for 1500 ms: "Correct" or "Incorrect" or "No response detected". Each block lasted approximately 7 min, and the overall experiment lasted approximately 35 min.

2.2. Results

To preview the results, the most important finding is the inverse distance effect for mixed comparisons (Fig. 1a). People were faster when mixed comparisons were near (−1 vs. 2) than when they were far (−1 vs. 7). This is inconsistent with both symbol+ and analog+. Table 2 presents the descriptive and inferential statistics concerning the distance, size, semantic congruence, and SNARC effects for all relevant comparison types.

We trimmed responses outside the interval of 200–2000 ms (0.44% of the data), and responses more than three standard deviations from each participant's mean (1.84% of the data). The error rate was low (M = 1.99%, SEM = 1.56%), and there was no speed-accuracy tradeoff (r = 0.561, p < .001, when correlating the mean response time and error rate across the 16 cells of the factorial design). Therefore, all analyses used response times on correct trials.

2.2.1. Magnitude effects

The magnitude effects are of primary interest as they directly inform the nature of the mental representation of integers. We first analyzed the distance effect for all four comparisons types. We then investigated the size effect for positive and negative comparisons.

2.2.1.1. Distance effect. Comparison type (positive, negative, mixed, zero), distance (near, far), and predicate (greater, lesser) were crossed in a repeated measures MANOVA. This section reports the distance (magnitude) effects. The predicate (judgment) effects are reported below.

There was a main effect of comparison type ($F(3,18)$ = 68.485, p < .001). The positive comparison, which is the standard in the literature, was designated an *a priori* contrast for the other comparison types. Positive comparisons were faster than negative comparisons ($F(1,20)$ = 155.818, p < .001), slower than mixed comparisons ($F(1,20)$ = 11.394, p = .003), and the same speed as zero comparisons (p = .599). There was also an overall main effect of distance ($F(1,20)$ = 11.441, p = .003), with far comparisons (M = 627 ms, SEM = 18 ms) faster than near comparisons (M = 639, SEM = 19).

Of greater interest is the reliable interaction between comparison type and distance ($F(3,18)$ = 15.528, p < .001). Positive and negative comparisons had comparable distance effects (p = .441). By contrast, positive and mixed comparisons had different distance effects ($F(1,20)$ = 37.101, p < .001), as did positive and zero comparisons ($F(1,20)$ = 6.880, p = .016).

Critically, mixed comparisons exhibited an inverse distance effect. Near comparisons (−1 vs. 2) were faster than far comparisons (−1 vs. 7). This pattern was stable across participants. Nineteen of 21 participants showed an inverse distance effect for mixed comparisons (p < .001, binomial test), which compares favorably with the 18 who showed a distance effect for positive comparisons (p = .001), and the 17 who showed a distance effect for negative comparisons (p = .007). The inverse distance effect is inconsistent with both symbol+ and analog+.

Given the omnibus effects, we conducted a more precise analysis of the mixed comparisons by partitioning them into mixed-positive and mixed-negative comparisons based on whether the positive number had a greater absolute value than the negative number (−1 vs. 7) or not (−7 vs. 1). A repeated measures MANOVA crossed comparison type (mixed-positive, mixed-negative) and distance (near, far). Mixed-positive comparisons were faster overall than mixed-negative comparisons ($F(1,20)$ = 4.584, p = .045). There was also a reliable inverse distance effect ($F(1,20)$ = 21.532, p < .001). Critically, there was no interaction (p = .748), indicating that the inverse distance effect was comparable for both comparison types.

There was no effect of distance for zero comparisons. We partitioned them into zero-positive comparisons (0

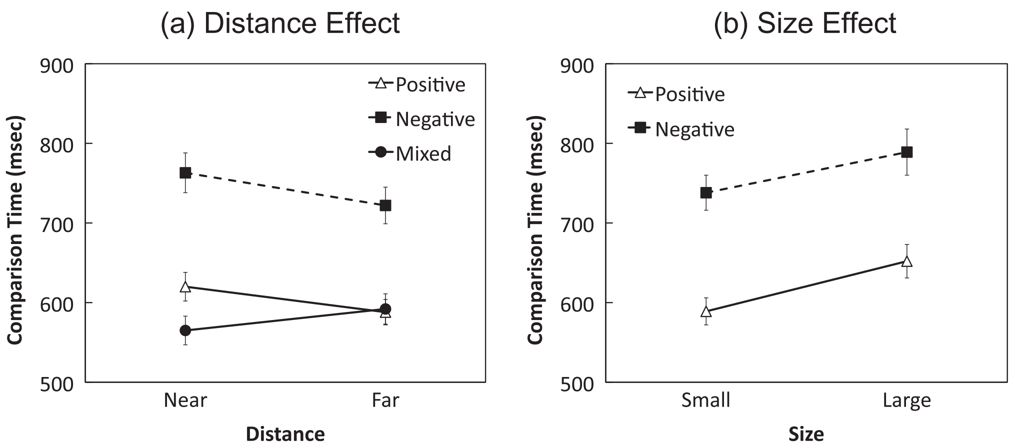


Fig. 1. Experiment 1 magnitude effects. Distance effect (a) for positive, negative, and mixed comparisons. (b) Size effect for positive and negative comparisons.

Table 2
Distance, size, semantic congruence, and SNARC effects for Experiment 1.

Distance effect	Overall[a]	Far	Near	t(20)[b]	p	Direction
Positive	604 (17)	588 (16)	620 (18)	4.324	<.001	Standard
Negative	742 (24)	722 (23)	763 (25)	4.464	<.001	Standard
Mixed	579 (18)	592 (19)	565 (18)	−4.917	<.001	Inverse
Mixed-positive	570 (17)	583 (19)	557 (17)	−2.456	.023	Inverse
Mixed-negative	585 (19)	600 (20)	570 (19)	−3.826	.001	Inverse
Zero	608 (17)	608 (17)	607 (18)	−0.072	.943	
Zero-positive	599 (18)	591 (18)	607 (20)	1.245	.227	
Zero-negative	614 (18)	623 (19)	606 (18)	−1.752	.095	(Inverse)
Size effect		**Near/small**	**Near/large**	**t(20)**	**p**	**Direction**
Positive		589 (17)	652 (21)	6.180	<.001	Standard
Negative		738 (22)	789 (29)	4.337	<.001	Standard
Semantic congruence effect		**Greater**	**Lesser**	**t(20)**	**p**	**Direction**
Positive		565 (17)	642 (19)	5.975	<.001	Standard
Negative		785 (27)	700 (22)	−5.178	<.001	Inverse
Mixed		580 (17)	577 (22)	−0.228	.822	
Mixed-positive		551 (17)	594 (21)	2.796	.011	Standard
Mixed-negative		600 (18)	565 (23)	−2.325	.031	Inverse
Zero		604 (17)	611 (21)	0.376	.711	
Zero-positive		595 (20)	605 (21)	0.470	.644	
Zero-negative		612 (19)	617 (23)	0.217	.831	
SNARC effect		**Slope[c]**		**t(20)**	**p**	**Direction**
Positive		0.95 (3.61)		0.263	.795	
Negative		1.10 (5.62)		0.195	.848	
Positive and negative		−0.81 (1.33)		−0.609	.550	

[a] Measures are reported *M* (*SEM*).
[b] Comparisons of near–far, large–small, lesser–greater, slope vs. 0.
[c] Predicting RT (right hand–left hand) from the mean of each near-distance number pair. A negative slope indicates a standard SNARC effect and a positive slope an inverse SNARC effect.

vs. 3) and zero-negative comparisons (−3 vs. 0) and conducted a repeated measures MANOVA crossing comparison type and distance (near, far). There were no main effects ($ps > 0.19$), but the interaction was reliable ($F(1,20) = 6.636$, $p = .018$). There was a marginal inverse distance effect for zero-negative comparisons, and no effect of distance for zero-positive comparisons (Table 2).

2.2.1.2. Size effect. The preceding analyses indicated that the distance effects were consistent with a magnitude representation of negative numbers. We next investigate whether this representation is psychophysically scaled by analyzing the size effect. The size effect can only be analyzed for near-distance comparisons (1 vs. 3, 7 vs. 9), because far-distance comparisons (1 vs. 9) include both small and large numbers. Recall that in these analyses, size is defined by the absolute value of the numbers being compared (e.g., −9 vs. −7 is a large comparison). A repeated measures MANOVA crossed comparison type (positive, negative), size (small, large), and predicate (greater, lesser). (Predicate effects are described with the judgment effects below.) The results appear in Fig. 1b and Table 2. There was a main effect of size ($F(1,20) = 46.494$, $p < .001$), with small comparisons ($M = 663$, $SEM = 19$) faster than large comparisons ($M = 721$, $SEM = 24$). There was a main effect of comparison type ($F(1,20) = 135.076$, $p < .001$), with positive comparisons faster than negative comparisons. The size effects for positive and negative comparisons were comparable, as indicated by the lack of an interaction ($p = .380$). This indicates that the negative numbers, like the natural numbers, have greater resolution when they are closer to zero.

2.2.2. Judgment effects

2.2.2.1. Semantic congruence effect. We report effects of the predicate factor (greater, lesser) from the preceding MANOVA of comparison type by distance by predicate. Fig. 2a shows the only reliable predicate effect: an interaction between predicate and comparison type ($F(3,18) = 25.850$, $p < .001$; other predicate effects, $p > .33$). The effect of predicate on positive comparisons was different from its effect on negative, mixed, and zero comparisons ($ps < .001$). Table 2 shows the results of paired *t*-tests evaluating the effect of predicate for each comparison type. Greater judgments were faster than lesser judgments for positive comparisons and mixed-positive comparisons – the standard semantic congruence effect. By contrast, lesser judgments were faster than greater judgments for negative and mixed-negative comparisons – an inverse semantic congruence effect. This suggests that people had associated lesser with negative numbers in this task. There was no effect of predicate for zero comparisons.

To examine mixed comparisons more closely, a repeated measures MANOVA crossed comparison type (mixed-positive, mixed-negative) and predicate (greater, lesser). Neither main effect was reliable ($ps > .13$), but as is evident in Fig. 2b, there was a reliable interaction ($F(1,20) = 24.112$, $p < .001$) such that mixed-positive

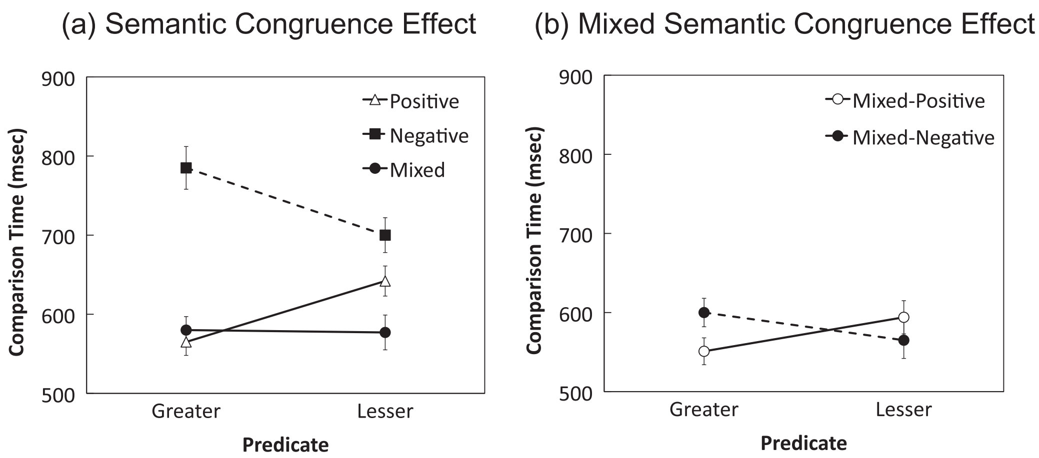


Fig. 2. Experiment 1 semantic congruence effect (a) for positive, negative, and mixed comparisons and (b) broken down for mixed-positive and mixed-negative comparisons.

comparisons showed the standard semantic congruence effect and mixed-negative comparisons showed the inverse effect (Table 2). A similar MANOVA on zero comparisons (zero-positive, zero-negative) crossed with predicate showed no effects ($ps > .20$).

The previous size effect analysis (small vs. large) included the predicate factor. We report the predicate effects here. There was no main effect of predicate ($p = .423$). There was a comparison type by predicate interaction ($F(1,20) = 45.815$, $p < .001$), consistent with the standard semantic congruence effect for positive comparisons and the inverse semantic congruence effect for negative comparisons. Critically, the three-way interaction of comparison type, size, and predicate was not reliable ($p = .449$). This implies that the semantic congruence effects are categorical: the association is between the predicate (greater and lesser) and sign of the numbers (positive and negative, respectively), and it is not graded by their magnitude.

2.2.2.2. SNARC effect. The design of the stimuli balanced correct responses across the right and left buttons, e.g., participants made greater judgments of both 1 vs. 3 (right response) and 3 vs. 1 (left response). A SNARC effect occurs when right responses are faster when judging pairs of large numbers and left responses are faster when judging pairs of small numbers. Therefore, the SNARC analysis only considers near-distance comparisons, because far-distance comparisons include both a small and large number. Three regressions were conducted for each participant: one for positive comparisons alone, one for negative comparisons alone, and one for positive and negative comparisons combined. For each regression, the independent variable was the mean of each number pair and the dependent variable was the mean right-hand response time minus the mean left-hand response time. A negative beta weight indicates a standard SNARC effect (left responses are faster for smaller number pairs and right responses are faster for larger number pairs), and a positive beta weight indicates an inverse SNARC effect (left responses are slower for smaller number pairs and right responses are slower for larger number pairs). Table 2 shows the average beta weight across participants for positive comparisons alone, negative comparisons alone, and positive and negative comparisons combined. None differed reliably from zero ($ps > .54$), indicating an absence of SNARC effects.

2.3. Discussion

Experiment 1 demonstrated standard distance and size effects for positive and negative comparisons, consistent with both symbol+ and analog+. Critically, mixed comparisons exhibited an inverse distance effect – near comparisons (-1 vs. 2) were faster than far comparisons (-1 vs. 7). This finding is inconsistent with symbol+, which proposes that people apply a rule for mixed comparisons ("positives are greater than negatives"), and therefore predicts no effect of distance. This finding is also inconsistent with analog+, which proposes that people consult a mental number line augmented with negative numbers and zero, and therefore predicts a standard distance effect. The inverse distance effect was stable across participants, and it held up when partitioning the date more finely into mixed-positive and mixed-negative comparisons.

The results indicate that number comparisons are facilitated when numbers closely straddle the zero boundary, as is the case for near-distance mixed comparisons. This is analogous to the finding of categorical perception whereby people are fast at discriminating sounds or colors that are physically close, but fall on different sides of a culturally important perceptual boundary (Harnad, 1990).

A number of studies have found evidence for the facilitating effects of number boundaries. Holyoak (1978) instructed participants to judge which of two numbers was closer to the boundary number 5. There was an inverse distance effect, with faster response times when the closer number was near the boundary (2 vs. 6) than when it was far (2 vs. 7). Other studies have found facilitating effects for place-value boundaries (e.g., Verguts & De Moor, 2005). Brysbaert (1995) found a boundary effect when participants compared pairs of two-digit numbers. When both numbers were on the same side of a decade boundary, there was a standard distance effect (61 vs. 69 was faster than 61 vs. 63). When the numbers crossed a decade boundary, there was an inverse distance effect (58 vs. 60 was faster than 52 vs. 60). To take another example, Franklin, Jonides, and Smith (2009) had participants judge whether triples of two-digit numbers were in ascending order or not. The triples varied in whether they crossed a tens boundary (18, 20, 21) or not (13, 15, 16), and also in the distance between the smallest and largest members. They found an inverse distance effect when the triples crossed a tens boundary, with near-distance triples judged faster than far-distance triples.

The current experiment did not find SNARC effects for positive or negative comparisons. Prior studies using integers have found a SNARC effect for positive comparisons (Fischer & Rottman, 2005; Shaki & Petrusic, 2005). However, these studies did not include mixed comparisons or zero comparisons, and this may have enabled participants to adopt task-dependent strategies. Also, SNARC effects are typically attenuated in more mathematically experienced participants as would be the case with the graduate students in this experiment (Dehaene et al., 1993; Fischer & Rottman, 2005).

3. Analog-*x*: the reflection model

The analog-x hypothesis proposes that the magnitude representation of number has been restructured by the integer symbol system. It is a 'computational theory' in Marr's (1982) analysis of cognitive theories: an abstract specification that must be instantiated at the middle level, by specifying the 'data structures' and 'algorithms' that implement the restructuring claim. Here, we develop one instantiation and show that it accounts for the results of Experiment 1, including the critical inverse distance effect for mixed comparisons, which falsifies the symbol+ and analog+ hypotheses. The *reflection model* implements analog-*x* using a spatial representation and continuous processing (i.e., distance computations) defined on this representation.[3] Integers are understood as points on mental number lines, and critically, the negative number line is a *reflection* of the natural number line. As a result of this restructuring, the integer number line differs from the conventional number line of mathematics.

This section develops the reflection model informally, with respect to Fig. 3. It focuses on the model's account of the magnitude (distance and size) effects of Experiment 1, and its qualitative implementation of the restructuring claim. The reader interested in further details of the model – a formal specification, a sensitivity analysis of its parameters, and an account of the judgment effects – is directed to Appendix A.

3.1. Model assumptions

The reflection model, shown in Fig. 3a, is defined by four assumptions.

1. Mental number lines are psychophysically scaled (with the degree of non-linearity specified by the α parameter).
2. The negative number line is a reflection and compression of the natural number line (with the degree of compression specified by the β parameter).
3. Comparing numbers requires projecting the corresponding points onto an orthogonal axis to compute the mathematical distance between them.
4. Comparison time is inversely proportional to mathematical distance.

Assumption (1) is inherited from prior research on mental number lines (Banks & Hill, 1974; Curtis & Fox, 1969; Dehaene & Mehler, 1992). Assumptions (3) and (4) are inherited from prior research on making decisions about numbers (Moyer & Landauer, 1967; Shepard, Kilpatric, & Cunningham, 1975). Assumption (2) is the critical new addition. It specifies how the natural number line is restructured to form the negative number line. (See Appendix A.1 for a formalization of these assumptions.)

3.2. Empirical evaluation

We evaluated the reflection model against the results of Experiment 1, using grid search to estimate the values of the two free parameters (α and β) that maximized the correlation between human and model comparison times computed across all stimuli. The model provides a good quantitative account of the data ($r = 0.94$, $p < .001$). (See Appendix A.2 for a sensitivity analysis indicating that a reduced model with just the β parameter fits the data as well as the full model.)

The reflection model also provides a good qualitative account of the distance and size effects found in Experiment 1 (Fig. 3a). It predicts that negative comparisons are slower than positive comparisons because the negative number line is compressed, and therefore negative magnitudes are less discriminable. It predicts that mixed comparisons are faster than positive comparisons because distances are greater when comparing across number lines than within a number line. The model predicts distance and size effects for positive comparisons and negative comparisons because the natural and negative number lines are psychophysically scaled. Critically, it predicts an inverse distance effect for mixed comparisons because of the reflective organization of the natural and negative number lines. Finally, the model predicts an inverse distance effect for zero-negative comparisons, again because of the reflective organization of the number lines.

The reflection model makes two incorrect predictions regarding the magnitude effects, both involving zero. First, it predicts that zero comparisons are faster than positive comparisons (because zero is the most extreme point of the natural number line, and therefore comparisons to it involve greater distances), whereas Experiment 1 found no difference between these comparison types. Second, it predicts a standard distance effect for zero-positive comparisons (because zero and positive numbers belong to the same number line), which was not observed in Experiment 1, although the trend was in the right direction. The model is not alone in these failings; numerous studies have found that zero behaves differently from other numbers (Brysbaert, 1995; Dehaene & Mehler, 1992; Nuerk, Iversen, & Willmes, 2004; Parkman, 1971). We return to the mental representation of zero below, in Section 6. (See Appendix

[3] Below, we instantiate the analog-x computational theory in a second model, one that adopts qualitatively different data structures and algorithms at the middle level.

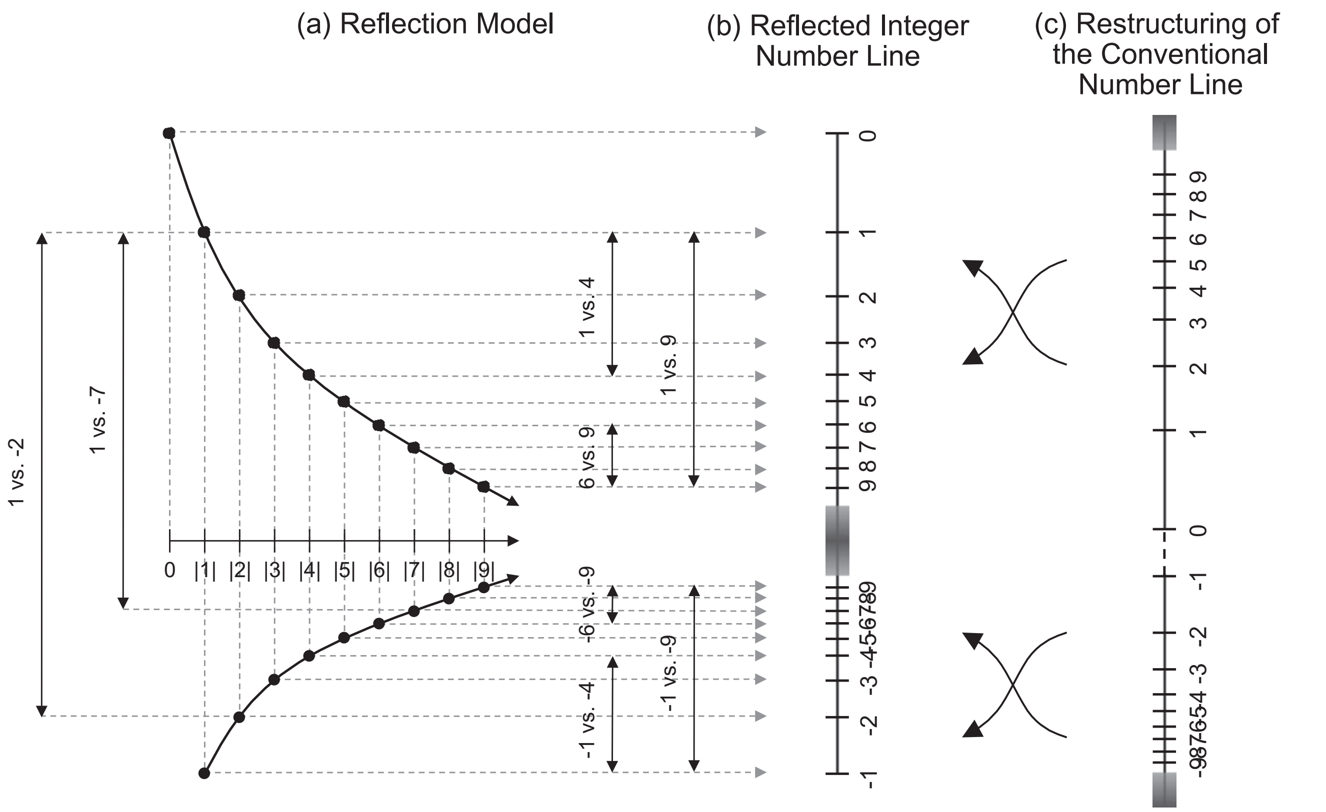


Fig. 3. (a) Reflection model with example distance and size effects. Points above the reference axis correspond to non-negative numbers and points below to negative numbers. (b) Reflected integer number line. (c) The reflected integer number line as a restructuring of the conventional mathematical number line.

A.3 for the model's account of the semantic congruence effect.)

3.3. Restructuring of spatial representations

The analog-*x* hypothesis proposes that mastering a new symbol system restructures existing representations to encode the unique properties of the symbol system. The reflection model defines an integer representation that includes more structure than the natural number representation, structure that is unique to the integer symbol system. It implies a reflected integer number line, as can be seen by projecting all points (corresponding to all numbers) onto a single dimension (Fig. 3b). Topologically, the reflected integer number line can be viewed as the product of "cutting" the conventional integer number line just below zero and "inverting" the natural and negative number lines (Fig. 3c). This psychological ordering of the integers differs from the conventional mathematical ordering in that the negative numbers reside beyond $+\infty$.[4]

One unique property of the integer symbol system is that zero is a boundary between positive and negative numbers. The reflection model encodes the boundary spatially by adjoining the negative number line to the $+\infty$ pole of the natural number line rather than the zero pole. This has the effect of magnifying the distance between negative and positive numbers close to zero. Another unique property is the additive inverse axiom: For every integer x, there exists an integer $-x$ such that $x + (-x) = 0$.[5] The model also encodes this property spatially, using symmetry. Integers are vertically aligned with their additive inverses, as shown in Fig. 3a.

The reflection model implements the analog-*x* hypothesis using a spatial representation, and encodes the unique properties of the integer symbol system using spatial transformations such as cutting and reflecting. The model makes the novel prediction that people incorporate symmetry into their integer representations, and should therefore be sensitive to symmetry when performing tasks other than comparison. We return to this prediction below, in Section 6.

4. Experiment 2

The analog-*x* hypothesis and its instantiation in the reflection model critically depend on the inverse distance effect found for mixed comparisons in Experiment 1 (cf. Krajcsi & Igács, 2010). However, Tzelgov et al. (2009) found no effect of distance for mixed comparisons. Given the discrepant findings, it is important to replicate the inverse

[4] Interestingly, the reflection model's placement of the negative numbers beyond $+\infty$ is consistent with Euler's (1755/2000) proposal "that negative numbers might sometimes be considered… more than infinity" (p. 57), i.e., that negative numbers are greater than positive numbers.

[5] This is the critical difference between the natural numbers and the integers. The natural numbers coupled with the addition operation satisfy the associativity, identity, and commutativity axioms, and therefore have the structure of a *commutative monoid*. The integers coupled with the addition operation additionally satisfy the inverse axiom, and therefore have the structure of a *commutative group*.

effect and generalize it to a different population and different experimental design. Experiment 2 repeated Experiment 1 with two main differences. First, the participants were drawn from a 2-year community college instead of graduate school. Second, the distance factor varied parametrically instead of being blocked into near and far levels.

4.1. Method

4.1.1. Participants

Fifty-five community college students (25 female; 5 left-handed) participated in the experiment for course credit. All reported normal or corrected-to-normal vision.

4.1.2. Stimuli and design

As in Experiment 1, there were three within-subjects factors: comparison type, distance, and predicate. The major difference was that comparison distance was varied parametrically. We constructed a matrix containing all possible comparisons of one-digit integers with distances ranging from 2 to 8. Because time constraints prohibited participants from making all possible comparisons, we randomly sampled 52 comparisons from the matrix for each participant, subject to several constraints. Each of the seven distances was sampled twice for each of the four comparison types, with two exceptions: For distance 8, there is only one positive comparison (1 vs. 9) and one negative comparison (-9 vs. -1), and therefore these cells of the matrix contained only one comparison each. For distance 2, there is only one mixed comparison (1 vs. -1), and it is symmetric, so this cell of the matrix was left empty. Another sampling constraint was that at each distance, the mixed comparisons included one mixed-positive comparison and one mixed-negative comparison. Similarly, at each distance, the zero comparisons included one zero-positive comparison and one zero-negative comparison. The sampling design covers a range of distances across the different comparison types, but it provides too few data points for evaluating the size effect. This trade-off was deemed acceptable because size effects for positive and negative comparisons have been consistently documented (Experiment 1; Ganor-Stern & Tzelgov, 2008; Shaki & Petrusic, 2005).

There were four experimental blocks. Each sampled comparison appeared once in each experimental block. Trials were randomly ordered within each block. The left–right order and greater-lesser judgment were counter-balanced across blocks. Participants were randomly assigned to one of two predicate orders (> < > < and < >< >).

4.1.3. Procedure

The experiment was implemented in E-Prime running on laptops with LCD displays measuring 35.75 cm diagonally. Responses were collected by pressing the "C" and "M" keys. Participants completed practice blocks of eight greater judgments and eight lesser judgments. Participants then completed the four experimental blocks. Each block began with seven unmarked practice trials before the recorded trials began.

Stimulus presentation was the same as Experiment 1 with the exception that there was no feedback. (This was done in preparation for an fMRI experiment that did not provide feedback.) Instead, a blank screen was presented between trials for a randomly determined interval of 2000, 4000, or 6000 ms. Each block lasted approximately 7 min, and the overall experiment lasted approximately 35 min.

4.2. Results

Data were lost from one participant due to a computer error. Another six participants were excluded because they confused whether to make greater or lesser judgments on at least one block. (Because there was no feedback, they could not self-correct.) Response times outside the range of 200–2000 ms were pruned (0.25%), as were response times more than three standard deviations from each participant's mean (1.49%). The average error rate was low ($M = 2.08\%$, $SEM = 2.02\%$), and there was no speed-accuracy trade off ($r = 0.549$, $p < .001$, when correlating response time and error rate across the 54 cells of the design).[6] All analyses therefore used response times on correct trials.

To preview the critical result, Fig. 4a shows that participants replicated the inverse distance effect for mixed comparisons. Table 3 shows the distance, semantic congruence, and SNARC effects for each comparison type.

4.2.1. Magnitude effects (distance effect)

A repeated measures MANOVA crossed comparison type (positive, negative, mixed, zero) and predicate (greater, lesser). There was a main effect of comparison type ($F(3,45) = 132.686$, $p < .001$). Positive comparisons were faster than negative comparisons ($F(1,47) = 291.715$, $p < .001$), slower than mixed comparisons ($F(1,47) = 7.939$, $p = .007$), and the same as zero comparisons ($p = .297$). The effect of predicate is presented below when considering the semantic congruence effect.

We used a regression approach to evaluate the distance effects (Fig. 4b). The slope (beta weight) of regressing response time on distance was computed for each comparison type for each participant. Participants' average slopes are shown in Table 3, which also holds the results of t-tests comparing the average slopes against the null hypothesis of a zero slope. There was a standard distance effect (negative slope) for positive comparisons and for negative comparisons. These slopes were comparable (paired $t(47) = 1.438$, $p = .157$). There was a reliable inverse distance effect (positive slope) for mixed comparisons. This result held at the individual level, albeit marginally. Thirty-one of 48 participants exhibited an inverse distance effect for mixed comparisons ($p = .059$, binomial test). By comparison, 37 participants showed a standard distance effect for positive comparisons ($p < .001$), and 36 showed a standard distance effect for negative comparisons ($p = .001$). When partitioned into mixed-positive and mixed-negative comparisons, there was an inverse distance effect for mixed-negative comparisons, but no effect of distance for mixed-positive comparisons (Table 3).

[6] For positive, negative, and zero comparisons: 7 distances × 2 predicates = 14 cells each. For mixed comparisons, which included no comparisons of distance 2: 6 distances × 2 predicates = 12 cells.

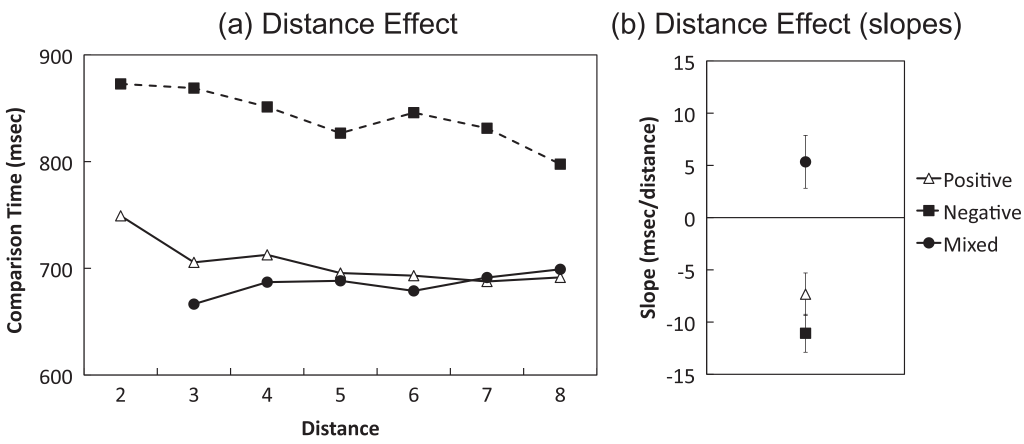


Fig. 4. Experiment 2 distance effect for positive, negative, and mixed comparisons expressed as (a) RTs and (b) slopes.

Table 3
Distance, size, semantic congruence, and SNARC effects for Experiment 2.

Distance effect	Overall[a]	Slope[b]	$t(47)$[c]	p	Direction
Positive	697 (15)	−7.33 (2.03)	−3.596	.001	Standard
Negative	836 (18)	−11.07 (1.82)	−6.084	<.001	Standard
Mixed	676 (16)	5.34 (2.53)	2.110	.040	Inverse
Mixed-positive	682 (19)	2.22 (3.44)	0.646	.522	
Mixed-negative	686 (17)	8.83 (3.10)	2.840	.007	Inverse
Zero	691 (16)	−1.43 (1.49)	−0.964	.340	
Zero-positive	692 (18)	−3.60 (1.93)	−1.864	.069	(Standard)
Zero-negative	708 (19)	0.35 (1.92)	0.186	.853	
Semantic congruence effect	**Greater**	**Lesser**	**$t(47)$**	**p**	**Direction**
Negative	866 (19)	805 (20)	−4.743	<.001	Inverse
Mixed	686 (18)	667 (16)	−1.738	.089	
Mixed-positive	671 (19)	679 (18)	0.592	.557	
Mixed-negative	700 (18)	654 (15)	−3.535	.001	Inverse
Zero	689 (16)	692 (17)	0.316	.753	
Zero-positive	673 (17)	693 (16)	1.625	.105	
Zero-negative	706 (17)	694 (20)	−0.737	.465	
SNARC effect	**Slope[d]**		**$t(47)$**	**p**	**Direction**
Positive	7.63 (6.05)		1.262	.213	
Negative	15.57 (7.29)		2.134	.038	Inverse
Positive and negative	0.633 (1.57)		0.402	.689	

[a] Measures are reported *M* (*SEM*).
[b] Average slope predicting RT from distance.
[c] Comparisons of slope vs. 0, lesser–greater.
[d] Predicting RT (right hand–left hand) from the mean of each near-distance number pair. A negative slope indicates a standard SNARC effect and a positive slope an inverse SNARC effect.

There was no distance effect for zero comparisons. When zero comparisons were partitioned based on the sign of the non-zero number, there was a marginal distance effect for zero-positive comparisons, but no effect of distance for zero-negative comparisons. This was somewhat different than the pattern in Experiment 1, where there was no effect for zero-positive comparisons, but a marginal inverse distance effect for zero-negative comparisons.

4.2.2. Judgment effects

4.2.2.1. Semantic congruence effect. Returning to the MANOVA of comparison type by predicate, there was no main effect of predicate ($p = .559$). Fig. 5a shows the reliable predicate by comparison type interaction ($F(1,45) = 26.912$, $p < .001$). Table 3 breaks down the interaction in a series of paired *t*-tests evaluating the effect of predicate for each comparison type and subtype. There was a semantic congruence effect for positive comparisons and an inverse semantic congruence effect for negative comparisons.

Fig. 5b shows the semantic congruence effects for mixed comparisons. A repeated measures MANOVA crossed comparison type (mixed-positive, mixed-negative) and predicate (greater, lesser). Neither main effect was reliable ($ps > .08$), but as in Experiment 1 there was an interaction ($F(1,47) = 17.931$, $p < .001$). Breaking down the interaction with paired *t*-tests, there was an inverse semantic congruence effect for mixed-negative comparisons, but no effect of semantic congruence effect for mixed-positive comparisons (Table 3).

For zero-comparisons, a repeated measures MANOVA crossed comparison type (zero-positive, zero-negative)

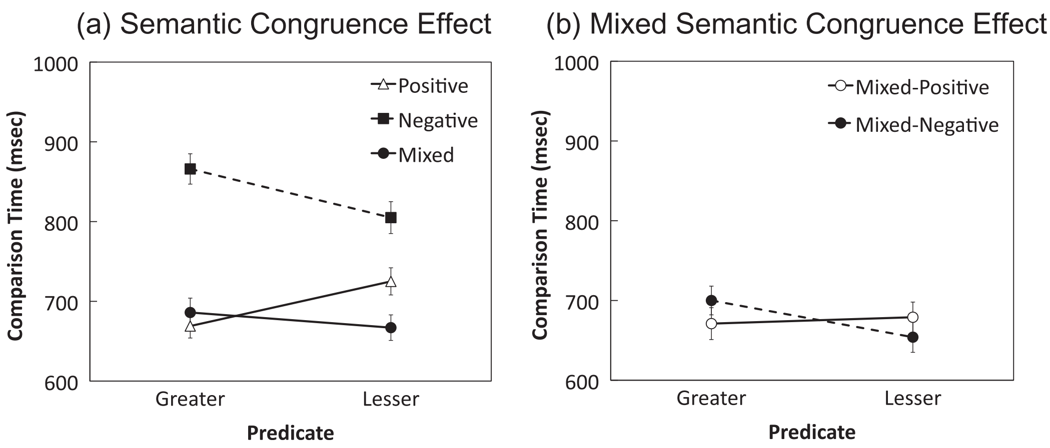


Fig. 5. Experiment 2 semantic congruence effect (a) for positive, negative, and mixed comparisons and (b) broken down for mixed-positive and mixed-negative comparisons.

and predicate (greater, lesser). As with Experiment 1, there was no main effect of predicate nor an interaction ($ps > .10$). There was a main effect of comparison type ($F(1,47) = 5.310$, $p = .026$), with zero-positive comparisons faster than zero-negative comparisons.

4.2.2.2. SNARC effect. Analyses of the SNARC effect focused on positive and negative comparisons of near distance (2–4), because the effect requires both numbers to be about same size. Three SNARC effects were estimated for each participant: one for positive comparisons alone, one for negative comparisons alone, and one for positive and negative comparisons combined. Table 3 shows the results. There was no SNARC effect for positive comparisons alone. There was an inverse SNARC effect for negative comparisons alone, with faster right responses for small negative numbers (−9 vs. −7) and faster left responses for large negative numbers (−3 vs. −1), replicating Fischer and Rottman (2005). There was no unified SNARC for both comparison types combined. The finding of an inverse SNARC effect for negative comparisons in Experiment 2 contrasts with the absence of SNARC effects in Experiment 1. We return to this in the Discussion.

4.3. Discussion

Experiment 2 replicated the inverse distance effect for mixed comparisons observed in Experiment 1 while sampling from a less-schooled population and using a parametric design. This finding is consistent with analog-*x* and inconsistent with both symbol+ (which predicts no effect of distance) and analog+ (which predicts a standard distance effect).

We can quantify the fit of the reflection model as in Experiment 1. The correlation between model and human comparison times was $r(92) = 0.91$, $p < .001$. (See Appendix A.2 for a sensitivity analysis showing that a reduced model with just the β parameter fits the data as well as the full model.) The model again provides a good qualitative account of the magnitude effects. It correctly predicts that positive comparisons are faster than negative comparisons and slower than mixed comparisons; that positive comparisons and negative comparisons show distance and size effects; and that mixed comparisons show an inverse distance effect. The model's incorrect predictions again concern zero comparisons. The model predicts that zero comparisons are faster than positive comparisons, but Experiment 2 found no difference. The model also predicts that zero-negative comparisons should show an inverse distance effect whereas Experiment 2 found no effect. These mispredictions, which parallel those of Experiment 1, represent limitations of the reflection model. We return to them below, in Section 6.

Experiments 1 and 2 found evidence that adults have access to an analog-*x* representation. Experiment 3 evaluated the developmental claim of the restructuring hypothesis: Children initially adopt a symbol+ representation, complementing their existing magnitude representation of natural numbers with the symbolic rules of the new integer symbol system. Over time, repeated application of the rules transforms the magnitude representation to encode the unique properties of the symbol system. The participants in Experiment 3 were children who had learned the integers, but had relatively little experience with them. The prediction is that they should rely exclusively on a symbol+ representation, and show no evidence of an analog-*x* (or analog+) representation.

5. Experiment 3

Children in the United States are introduced to negative numbers around 4th grade. At this point, they already have a magnitude representation of natural numbers (Duncan & McFarland, 1980; Sekuler & Mierkiewicz, 1977). They learn about negative numbers through informal semantic models (Schwarz, Kohn, & Resnick, 1993/1994). In "annihilation" semantics, they are told that natural numbers and negative numbers can be represented as sets of positively and negatively charged particles, respectively, and there is a rule for canceling particles of opposite charges (Hayes & Stacey, 1999; Liebeck, 1990). In "extended number line" semantics, they are told that the number line can be extended to the left of zero, and that natural numbers correspond to rightward movements and negative numbers to leftward movements (Hativa & Cohen, 1995; Moreno & Mayer, 1999).[7]

Informal semantic models generalize poorly to arithmetic operations besides addition, such as subtracting a negative number from another integer (Hayes & Stacey, 1999; Liebeck, 1990; Moreno & Mayer, 1999). For this reason, instruction quickly transitions to the rules of the integer symbol system (Thompson & Dreyfus, 1988). Children learn rules for comparing integers such as "positives are greater than negatives," and rules for integer arithmetic such as "a negative times a negative is a positive."[8]

Experiment 3 investigates whether sixth-grade children (12 year olds), who have had a couple of years of experience with negative numbers, use a symbol+, analog+, or analog-*x* representation. The restructuring hypothesis predicts that they have not had sufficient experience to restructure their magnitude representation of natural numbers, and will therefore use a symbol+ representation (Prather & Alibali, 2008). In particular, when making mixed comparisons, they will invoke the rule "positives are greater than negatives," and will therefore show no effect of distance. Similarly, when making zero comparisons, they will invoke rules such as "positives are greater than zero," and will therefore not show an effect of distance.

5.1. Method

5.1.1. Participants

Thirty-six children (M = 11.89 years, SD = 0.34; 17 female; 3 left-handed) participated. They were in an accelerated sixth-grade mathematics course covering the first half of pre-algebra. All reported normal or corrected-to-normal vision.

5.1.2. Stimuli and design

The design was similar to Experiment 1 with two changes. To reduce participant confusion, the predicate factor was between-subjects: Half the participants made only greater judgments, and half made only lesser judgments. The second change was to reduce the number of stimuli by half. There were four experimental blocks, each containing 44 stimuli. Each comparison appeared once in each block, counter-balanced across blocks for left–right order.

5.1.3. Procedure

The apparatus was the same as in Experiment 2. The procedure was the same as Experiment 1 with the exception that participants only completed one practice block with their assigned predicate (greater or lesser), and experimental blocks did not begin with unmarked practice trials. Each block lasted approximately 4 min, and the overall experiment lasted approximately 20 min, fitting within the constraints of the school day.

5.2. Results

Fig. 6a previews the main result. There was no effect of distance for mixed comparisons, corroborating the developmental prediction of the restructuring hypothesis.

Responses outside the range of 200–2000 ms were pruned (0.21%), as were response times more than three standard deviations from each participant's mean (1.42%). There was no speed-accuracy trade-off (r = 0.763, p = .001, between response time and error rate across the 16 cells of the design). The error rate was low (M = 3.79%, SEM = 1.91%). Analyses therefore focused on response times on correct trials.

5.2.1. Magnitude effects

5.2.1.1. Distance effect. We performed a MANOVA on repeated measures comparison type (positive, negative, mixed, zero) and distance (near, far), and between-subjects factor predicate (greater, lesser). Table 4 shows the associated means and standard errors. Predicate effects are described below, in the section on judgment effects. There was a main effect of comparison type ($F(3,32) = 41.036$, $p < .001$), with positive comparisons faster than negative comparisons ($F(1,34) = 77.506$, $p < .001$), slower than mixed comparisons ($F(1,34) = 9.267$, $p = .004$), and slower than zero comparisons ($F(1,34) = 5.494$, $p = .025$). There was a main effect of distance ($F(1,34) = 14.483$, $p = .001$), with far comparisons (M = 647, SEM = 15) faster than near comparisons (M = 663, SEM = 15). There was an interaction between comparison type and distance ($F(3,32) = 4.802$, $p = .007$). Table 4 shows the results of paired *t*-tests breaking down the interaction. There was a distance effect for positive comparisons and one for negative comparisons, and they were comparable (p = .984).

Critically, there was no effect of distance for mixed comparisons, consistent with symbol+. A repeated measures MANOVA on comparison type (mixed-positive, mixed-negative) and distance (near, far) found no main effects of comparison type or distance ($ps > .24$), but there was a reliable interaction ($F(1,35) = 4.825$, $p = .035$). However, this interaction was not strong; paired *t*-tests found a marginal distance effect for mixed-negative comparisons, and no effect of distance for mixed-positive comparisons (Table 4).

There was no effect of distance for zero comparisons. A repeated measure MANOVA on comparison type (zero-positive, zero-negative) and distance (near, far) found no reliable effects ($ps > .22$).

5.2.1.2. Size effect. Fig. 6b and Table 4 show the size effects for positive and negative near-distance comparisons. A repeated measures MANOVA crossing comparison type (positive, negative) and size (small, large) found a main effect of comparison type ($F(1,35) = 29.764$, $p < .001$), with positive comparisons faster than negative comparisons. There was also a main effect of size ($F(1,35) = 69.107$, $p < .001$), with small comparisons faster than large comparisons. The size effect was comparable for both compar-

[7] Recent studies of "operational momentum" effects during arithmetic are delineating the connection between number processing on one hand and pointing movements (Pinhas & Fischer, 2008) and eye movements (Knops, Thirion, Hubbard, Michel, & Dehaene, 2009) on the other.

[8] Children are not alone in requiring symbol systems to move beyond informal semantics. As Gauss wrote in a letter to Bessell, "when the definition, from which we proceed, ceases to have a sense, one should not ask, strictly speaking, what has to be assumed?, but what is convenient to assume? so that I can always remain consistent. Thus, for example, the product of minus by minus" (cited on p. 256 of Crowley & Dunn, 1985).

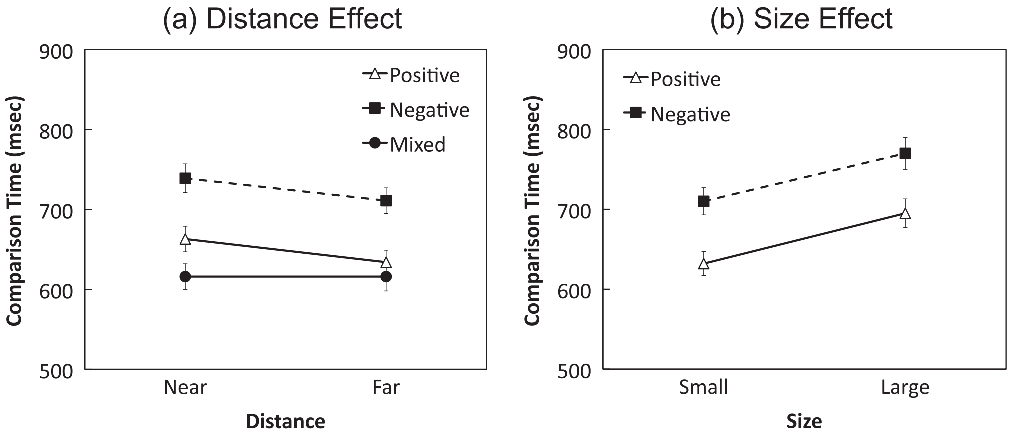


Fig. 6. Experiment 3 magnitude effects. (a) Distance effect for positive, negative, and mixed comparisons. (b) Size effect for positive and negative comparisons.

Table 4
Distance, size, semantic congruence, and SNARC effects for Experiment 3.

Distance effect	Overall[a]	Far	Near	$t(35)$[b]	p	Direction
Positive	649 (15)	634 (15)	663 (16)	4.376	<.001	Standard
Negative	725 (17)	711 (16)	739 (18)	3.848	<.001	Standard
Mixed	616 (17)	616 (18)	616 (16)	−0.048	.962	
Mixed-positive	619 (19)	627 (20)	612 (19)	−1.564	.127	
Mixed-negative	611 (16)	605 (17)	617 (15)	1.767	.086	(Standard)
Zero	631 (16)	627 (17)	634 (17)	0.875	.388	
Zero-positive	625 (16)	617 (17)	633 (17)	1.521	.137	
Zero-negative	637 (19)	638 (20)	636 (20)	−0.150	.882	
Size effect		Near/small	Near/large	$t(35)$	p	Direction
Positive		632 (15)	695 (18)	7.096	<.001	Standard
Negative		710 (17)	770 (20)	5.696	<.001	Standard
Semantic congruence effect		Greater	Lesser	$t(34)$	p	Direction
Positive		619 (22)	678 (22)	1.885	.068	(Standard)
Negative		753 (24)	698 (24)	−1.609	.117	
Mixed		635 (24)	596 (24)	−1.153	.257	
Mixed-positive		636 (27)	606 (27)	−0.765	.299	
Mixed-negative		635 (22)	589 (22)	−1.463	.304	
Zero		631 (23)	630 (23)	−0.021	.983	
Zero-positive		605 (23)	645 (23)	1.214	.233	
Zero-negative		655 (27)	619 (27)	−0.915	.367	
SNARC effect			Slope[c]	$t(35)$	p	Direction
Positive			−1.40 (3.42)	−0.410	.684	
Negative			3.59 (3.53)	1.018	.316	
Positive and negative			6.74 (1.87)	3.604	.001	Inverse

[a] Measures are reported *M* (*SEM*).
[b] Comparisons of near–far, large–small, lesser–greater, slope vs. 0.
[c] Predicting RT (right hand–left hand) from the mean of each near-distance number pair. A negative slope indicates a standard SNARC effect and a positive slope an inverse SNARC effect.

ison types as indicated by the absence of an interaction ($p = .763$).

5.2.2. Judgment effects

5.2.2.1. Semantic congruence effect. To evaluate the semantic congruence effect, we return to the distance effect MANOVA reported above. There was no main effect of predicate ($p = .780$). Fig. 7 shows the only reliable interaction, which was between comparison type and predicate ($F(3,32) = 15.311$, $p < .001$). Table 4 breaks down the interaction in a series of independent t-tests (recall that predicate is a between-subjects factor in this experiment) evaluating the effect of predicate for each comparison type and subtype. Importantly, the only semantic congruence effect was a marginal effect for positive comparisons, consistent with symbol+.

5.2.2.2. SNARC effect. SNARC effects were analyzed as in Experiment 1. Table 4 shows the results. There was no effect for positive comparisons alone nor for negative comparisons alone. However, there was an inverse SNARC effect across both comparison types combined, with left responses faster for positive comparisons and right responses faster for negative comparisons.

5.3. Discussion

Experiment 3 evaluated the prediction that children who have less experience with integers will not have restructured their magnitude representation, and will therefore understand negative numbers by using a combination of a natural number line representation and the rules of the integer symbol system. The results corroborate this prediction. The sixth graders showed distance effects for positive and negative comparisons. However, for the diagnostic mixed comparisons, they did not show the inverse distance effect predicted by analog-*x* and observed for adults in Experiments 1 and 2. Nor did they show the standard distance effect predicted by analog+. Instead, they showed no effect of distance, consistent with the use of rules ("positives are greater than negatives"), and thus with symbol+.

This null result, though predicted *a priori* by the restructuring hypothesis, must be interpreted with caution. It might, for example, simply reflect a lack of power. There are three reasons to believe that this is not the case. First, the semantic congruence effects also pattern according to symbol+, as predicted by the restructuring hypothesis. Symbol+ proposes that only positive numbers have magnitude representations. Consistent with this proposal, there was a marginal semantic congruence effect for positive comparisons, but no effect for negative and mixed comparisons. This suggests that the children lacked a separate magnitude representation of negative numbers to associate with the lesser predicate. The second reason comes from a direct comparison of the distance effects for mixed comparisons in Experiments 1 and 3, which used the same stimuli and similar designs. A 2 × 2 MANOVA crossing between-subjects factor age-group (adults, children) and within-subjects factor distance (far, near) found an age-group by distance interaction ($F(1,55) = 9.613$, $p = .003$), confirming that the effect of distance for mixed comparisons was statistically different for adults vs. children. Third, we have replicated the predicted null effect of distance for mixed comparisons observed in Experiment 3 in a separate study of 54 middle and secondary school students from a low-achieving urban school district in the United States (Varma, Harris, Schwartz, & Martin, 2009).

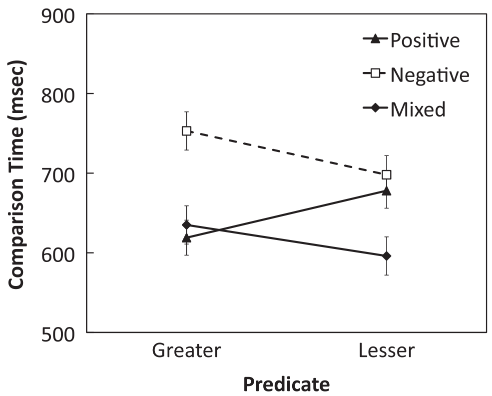


Fig. 7. Experiment 3 semantic congruence effect for positive, negative, and mixed comparisons.

(The other results of this study were also consistent with those of Experiment 3, including distance effects for positive and negative comparisons, and a semantic congruence effect for positive comparisons but none for negative and mixed comparisons.)

6. General discussion

As part of the larger question of how people come to understand abstract mathematical concepts, three experiments investigated how people represent the integers, including negative numbers and zero. The symbol+ hypothesis proposes that people supplement their magnitude representation of natural numbers with the rules of the integer symbol system. The analog+ hypothesis proposes that the mental number line is extended "to the left" to include negative numbers and zero *per* the conventional number line of mathematics. The existing literature is mixed, with some findings supporting symbol+, others analog+, and still others inconsistent with both hypotheses.

The current research investigated whether adults have additionally developed a restructured analog-*x* representation. The critical supporting finding was the inverse distance effect for mixed comparisons across the zero boundary, with near comparisons (−1 vs. 2) faster than far comparisons (−1 vs. 7). This effect was present whether distance was blocked or varied parametrically, and it was found in both highly educated graduate students and less educated junior college students. (The other magnitude and judgment effects in Experiments 1 and 2 were also consistent with analog-*x*.) Analog-*x* was instantiated in a formal model using spatial number line representations and continuous processing. The model encodes the additive inverse property of the integer symbol system by positing that the negative number line is a reflection of the positive number line. It provides a good quantitative and qualitative account of the adult data in Experiments 1 and 2, although it did not provide a good account of the representation of zero.

Experiment 3 evaluated the restructuring hypothesis, which makes a developmental claim: When first learning the integer symbol system, children initially adopt a symbol+ representation, but over time, repeated application of symbolic rules restructures the magnitude representation of natural numbers. As predicted, sixth-grade children did not show an effect of distance for mixed comparisons, even though they were as fast and accurate as adults. This finding is consistent with a symbol+ representation (as were the other observed magnitude and judgment effects).

6.1. The analog-x hypothesis

The analog-*x* hypothesis proposes that people come to understand abstract mathematical concepts by developing an analog representation of their important symbolic properties. In some cases, analog representations have a near one-to-one correspondence with perceptual experiences. For example, people represent the geometry of the rotation of physical objects much like they perceive the rotation it-self (Shepard & Metzler, 1971). For the case of natural numbers, people appear to utilize a mental analog of the perceptible number line (Moyer & Landauer, 1967). fMRI studies indicate that the intra-parietal sulcus (IPS) is a neural correlate of this mental number line. The IPS shows a "neural distance effect," with greater activation for difficult near comparisons than for easy far comparisons (Pinel, Dehaene, Rivière, & Le Bihan, 2001). Furthermore, fMRI studies suggest that the IPS represents number magnitudes and perceptual magnitudes similarly, because it also shows a neural distance effect when comparing the physical size and luminance of symbols (Pinel, Piazza, Le Bihan, & Dehaene, 2004) and the numerosity of dot clouds (Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004).

Analog-*x* differs from other accounts of analog representation in proposing that people can also develop analog representations that do not directly correspond to perceptual-motor experiences. This is critical for making sense of abstract mathematical concepts that do not have ready perceptual-motor referents. For the case of the integers, the reflection model proposes that people develop an analog-*x* representation where the positive and negative number lines are symmetrically organized. This symmetry captures the additive inverse property unique to the integer symbol system.

The primary evidence that people incorporate symmetry comes from the inverse distance effect for mixed comparisons. To gain further evidence for the symmetry proposal of analog-*x*, it is important to consider other experimental paradigms besides number comparison. Recent studies have directly examined the symmetry proposal using a number bisection task. Tsang and Schwartz (2009) showed adult participants pairs of integers and had them name the midpoint. They found evidence of a symmetry tuning curve: Participants were fastest for perfectly symmetric pairs (-3 and 3), and faster for more symmetric pairs (-7 and 9) than less symmetric pairs (-6 and 10).[9] In an fMRI study using the same task, Tsang, Rosenberg-Lee, Blair, Schwartz, and Menon (2010) found that more symmetric pairs elicited greater activation in left lateral occipital cortex, a region associated with visual symmetry processing (Sasaki, Vanduffel, Knutsen, Tyler, & Tootell, 2005). Thus, the new symmetry predictions generated by analog-*x* are finding empirical support.

6.2. The restructuring hypothesis

Taken together, Experiments 1 and 3 constituted a cross-sectional comparison. Their findings indicated that symbol+ precedes analog-*x* in development. Children used rules to judge that positives are greater than negatives, that positives are greater than zero, and so forth. They exhibited no magnitude effects when using these rules. By contrast, adults did not use rules, as evidenced by the finding of an inverse distance effect for mixed comparisons. This supports the restructuring hypothesis, which proposes that through experience using symbol systems (i.e., a symbol+ representation) people develop an analog-*x* representation that organizes positive and negative magnitudes symmetrically.

Developmental neuroscience studies have found independent evidence for a symbolic-to-analog shift in number understanding. Ansari, Garcia, Lucas, Hamon, and Dhital (2005) had 10-year-olds and adults compare natural numbers. Children showed a neural distance effect in inferior frontal gyrus, which is a neural correlate of controlled rule processing (Miller & Cohen, 2001). By contrast, adults showed a neural distance effect in IPS, which is associated with magnitude processing. Rivera, Reiss, Eckert, and Menon (2005) found evidence for a developmental shift using a symbolic arithmetic task. Through adolescence, activation in prefrontal regions associated with controlled processing decreased, whereas activation in posterior regions associated with visuospatial processing (i.e., IPS) increased. An interesting question for future research is whether IPS is the sole neural correlate of negative magnitudes, or whether it functions as part of a large-scale cortical network (Luria, 1966; Mesulam, 1990; Just & Varma, 2007).

6.3. Symbolic-magnitude model

Analog-*x* is a computational theory, specifying the abstract characteristics of integer understanding – symmetry and sensitivity to boundaries – but not their implementation at the level of data structures and algorithms (Marr, 1982). The reflection model instantiates analog-*x* using spatial representations (mental number lines) and continuous processing (distance computations). It offers a successful account of adult performance and generates novel theoretical predictions, for example regarding performance on bisection tasks.

Here, we propose a second instantiation of analog-*x* in a *symbolic-magnitude* model of integer understanding. The model serves two purposes. First, it offers an explanation of how symbol systems can restructure analog representations to incorporate additional structure. This is important because some researchers claim that such transformations are impossible because symbolic representations and analog magnitude representations are fundamentally incommensurable (e.g., Harnad, 1990). Second, the model offers a different implementation of analog-*x* at the level of data structures and algorithms, using featural representations of number magnitude (Lewis & Varma, 2010; Zorzi & Butterworth, 1999) and similarity-based processing.

This section describes the symbolic-magnitude model informally, with reference to Fig. 8. It evaluates its quantitative and qualitative fit to the results of Experiments 1 and 2, and puts forth a *concatenative* account of how symbolic rules restructure magnitude representations. The reader interested in a formalization of the model, a sensitivity analysis of its parameters, and its application to judgment effects is directed to Appendix B.

6.3.1. Model description

The symbolic-magnitude model for the integers -9 through 9 is shown in Fig. 8a. (See Appendix B.1 for the mathematical formalization.)

[9] Although accuracy data were not reported, there was no evidence of a speed-accuracy trade-off (J.M. Tsang, personal communication, September 28, 2010).

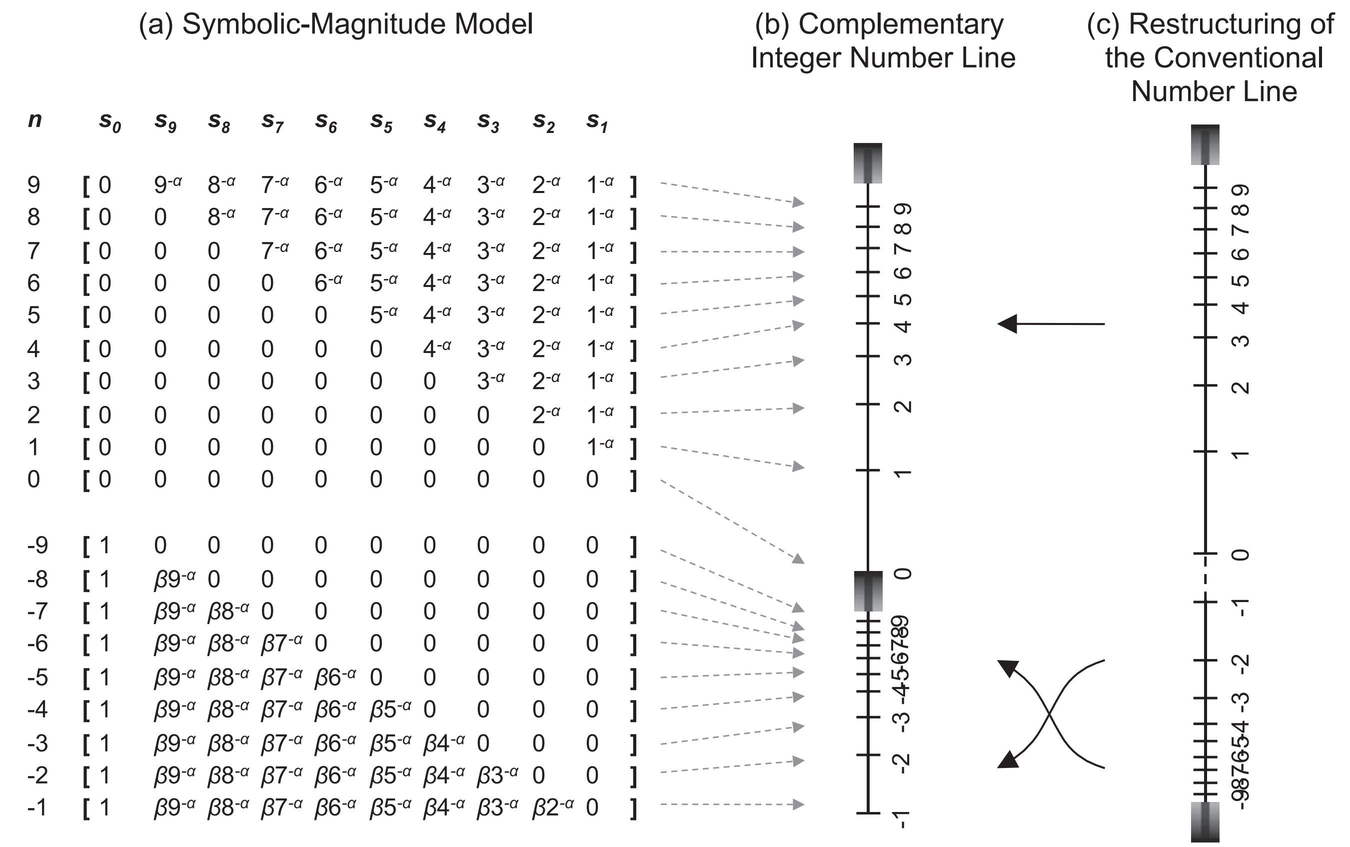


Fig. 8. (a) Symbolic-magnitude model. (b) Complementary integer number line. (c) Formation of the complementary integer number line via restructuring of the conventional integer number line.

Natural numbers are represented as vectors of features, as shown in the top half of Fig. 8a. Feature s_0 is the sign feature; its value is 0 for all natural numbers. Features s_1 through s_9 encode the magnitude of a number. Their values are psychophysically scaled, with the non-linearity of the scaling specified by the α parameter.

The critical proposal of analog-*x* is that negative number representations are symmetric to natural number representations. The reflection model implements this proposal using reflected number lines. By contrast, the symbolic-magnitude model defines the feature values of negative and natural number representations to be *complementary*, as shown in the bottom half of Fig. 8a. This is true of the sign feature s_0, which is 1 for negative numbers. This is also true of the magnitude features s_1 through s_9, where the magnitude encoding of a negative number is the complement of the magnitude encoding of the positive number that is its additive inverse.[10] For example, the encoding of -1 has a zero value for feature s_1 and non-zero values for all other features, whereas the encoding of 1, its additive inverse, has a non-zero value for feature s_1 and zero values for all other features. In addition, the feature values of negative magnitudes are compressed relative to those of positive magnitudes, as specified by the β parameter.

Numbers are compared using the similarity-based processing that is conventional for featural representations (Nosofsky, 1984; Shepard, 1987). This entails two steps. First, the difference between two number representations is computed by summing the differences of their feature values. Second, the similarity of the number representations is computed as a decreasing function of their difference: the less their difference, the greater their similarity. The time to compare the number representations is an increasing function of their similarity: the greater their similarity, the greater (i.e., slower) their comparison time.

Critically, the model predicts an inverse distance effect for mixed comparisons. For example, as Fig. 8a shows, the featural difference between the representations of -1 and 2 is greater than the featural difference between the representations of -1 and 7. Thus, the similarity of the near-distance pair is less than the similarity of the far-distance pair, and consequently the near-distance pair is compared faster than the far-distance pair.

6.3.2. Empirical evaluation

The symbolic-magnitude model provides a good quantitative account of the adult data. Using grid search to estimate the best-fitting values of the two free parameters (α and β), the correlation between model and human comparison times was $r = 0.96$ ($p < .001$) for Experiment 1 and $r = 0.92$ ($p < .001$) for Experiment 2. (See Appendix B.2 for a sensitivity analysis showing that a reduced model with just the β parameter yields comparable fits.)

[10] The complementary number representations of the symbolic-magnitude model are analogous to those of 2's complement, the dominant binary code for signed numbers.

The model also provides a good qualitative account of the distance and size effects observed in Experiments 1 and 2. It predicts that negative comparisons are slower than positive comparisons because the feature values of negative number representations are compressed (as specified by β), decreasing the featural difference between them, and thus increasing comparison time. It predicts that mixed comparisons are faster than positive comparisons because the feature values of positive and negative number representations are complementary, increasing the featural difference between them, and thus decreasing comparison time. It predicts that positive comparisons and zero comparisons are comparable because the zero representation is parametrically in line with positive number representations. The model predicts distance and size effects for positive comparisons and negative comparisons because of the psychophysical scaling of feature values. And, as described above, it predicts an inverse distance effect for mixed comparisons because the complementary coding of positive and negative numbers increases the featural difference between numbers close to the zero boundary, thus decreasing comparison time.

The symbolic-magnitude model provides a better qualitative account of the zero comparison data than the reflection model. It predicts a distance effect for zero-positive comparisons because natural number representations are psychophysically scaled. Supporting this prediction, there was a marginal distance effect in Experiment 2, and a trend for such an effect in Experiment 1. The model predicts an inverse distance effect for zero-negative comparisons because zero is a natural number and natural numbers and negative numbers are complementarily coded. Supporting this prediction, there was a marginal inverse distance effect in Experiment 2. The failure to find an effect of distance in Experiment 1 is inconsistent with both the model and the results of Experiment 2, and therefore difficult to interpret. (See Appendix B.3 for the model's account of the semantic congruence effect.)

6.3.3. Concatenative restructuring of featural representations

The symbolic-magnitude model specifies a restructured integer representation (Fig. 8b) that can be viewed as the product of cutting the conventional integer number line just below zero and inverting the negative number line (Fig. 8c). This psychological ordering of the integers represents a restructuring of the conventional mathematical ordering.

The model's featural representations provide a basis for understanding the developmental process by which a symbol+ representation is restructured into an analog-*x* representation. Children initially possess the natural number representations shown in the top half of Fig. 8a. Features s_1 through s_N are *interpreted* as a magnitude representation and processed via similarity-based processing. When learning the integer symbol system, children add the s_0 feature denoting the sign of a number. This feature is *interpreted* as a symbolic representation (e.g., the symbolic rule "positives are greater than negatives" can be operationalized as "if $s_0 = 0$ for x and $s_0 = 1$ for y, then $x > y$"). The result is a symbol+ representation, with features s_1 through s_N interpreted as a magnitude and feature s_0 interpreted symbolically.

Over development, children begin to articulate separate negative number representations, as shown in the bottom half of Fig. 8a. The symmetry of these representations follows from repeated application of the additive inverse property when solving informal problems (e.g., paying \$3 reduces a \$3 debt to \$0) and formal problems (e.g., eliminating terms in algebraic equations). Once children possess both natural number and negative number representations, the magnitude and symbolic features cease to be interpreted separately and come to be interpreted integrally, both processed through computing featural differences and representational similarities. (Of course, if the context calls for it, people can still interpret the s_0 feature symbolically, yielding behaviors consistent with symbol+.)

This concatenative account of the restructuring process is possible because the symbolic-magnitude model proposes a common number representation – vectors of feature values – throughout development. What changes is how this representation is interpreted. Initially, some features are processed via similarity and others symbolically, and the representation is therefore symbol+. Later, when all features have come to be processed via similarity, the representation becomes analog-*x*.

6.4. Future research

The current experiments and models suggest a number of directions for future research. Developmental studies are necessary to better evaluate the restructuring hypothesis. Cognitive neuroimaging studies are necessary to understand the neural bases of the analog-*x* representation in adults. Developmental neuroimaging studies are necessary to investigate the shifting cortical networks that underlie the proposed representational restructuring. In advancing our empirical understanding of the basic phenomena, these studies will also support evaluation of the predictions of the reflection and symbolic-magnitude models.

In this section, we sketch two more speculative directions for future research. The first is to better understand the type of experiences with integers that drive restructuring. We hypothesize that instruction that emphasizes the additive inverse property through experiences with symmetry about zero will accelerate the development of analog-*x* (Schwartz, Blair, & Tsang, in press). Such instruction is not currently the norm in mathematics classrooms, with most approaches emphasizing annihilation or counting forward and backward along the number line. A related question is whether accelerating the development of analog-*x* for integers better prepares people to learn new, but related, abstract mathematical concepts. One possibility is that analog-*x* might facilitate algebraic thinking, which requires symmetric canceling and equation balancing to maintain equality. Or conversely, it might be the symbolic experiences in algebra that drive the shift from symbol+ to analog-*x*.

A second goal for future research is to understand the generality of the abstract-to-concrete shift proposed here. The integers are one instance of an abstract mathematical concept. It is not difficult to imagine other instances, where initially abstract concepts become embodied in transformed analog magnitude representations. For example, "instantaneous acceleration" is an abstract physical concept. It is intuitive to physicists, but at one point in their training, it probably was not. A good deal of experience was likely necessary before the rate of change of a rate of change occurring in an interval of duration zero became intuitive. The restructuring hypothesis proposes that this was the result of sustained mathematical experiences, most likely with calculus. Through practice with mathematical symbol systems, budding physicists come to construct structured analog magnitude representations of movement that make abstract concepts like instantaneous acceleration concrete.

The restructuring hypothesis proposes connecting symbolic representations and processes with analog representations, so those analog representations can be transformed to encode the symbolic properties. Therefore, analog-*x* predicts that neither pure "symbol pushing" nor pure "hands-on" experiences will be sufficient for developing an analog-*x* representation. Determining the right balance is an important educational question which does not naturally fall out of the symbolic or embodied traditions of explaining mathematical competence.

6.5. Conclusion

It was said of the mathematician Ramanujan that "every positive integer was one of his personal friends" (Newman, 1956, p. 375). But mathematicians are not the only ones for whom mathematical concepts are alive – this is true to some degree for all who are mathematically literate. We have argued that repeated application of symbolic rules transforms magnitude representations, adding structure to them, and we have offered two models of the outcome. The transformed representations support an intuitive understanding of abstract mathematical concepts. They explain how mathematical symbol systems give ready access to the imperceptible.

Acknowledgements

We thank Robb Lindgren, Janet Go, Kristen Blair, and Jessica Tsang for comments on prior versions of this paper. This research was supported by the National Science Foundation under Grants REC 0337715, SLC 0354453, SGER 0742113, and REESE 07595. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Appendix A. The reflection model

This appendix formalizes the reflection model, quantifies its fit to the results of Experiments 1 and 2, presents a sensitivity analysis of its free parameters, and summarizes its account of judgment effects.

A.1. Model formalization

The first assumption of the reflection model is that mental number lines are psychophysically scaled. We formalize it by following the psychophysical theory of Stevens (1970), which specifies power function scaling: Psychological (number) magnitudes $m(x)$ are a power function of physical magnitudes x.

$$m(x) = (x+1)^{-\alpha} \quad x \geqslant 0 \tag{1}$$

The $\alpha \in (0,1]$ parameter specifies the curvature of mental number lines. Power function scaling of mental number lines has been proposed by a number of researchers (Banks & Hill, 1974; Curtis & Fox, 1969; Dehaene & Mehler, 1992).

The second assumption is that the negative magnitudes $s(x)$ are reflections and compressions of natural number magnitudes.

$$s(x) = \begin{cases} m(x) & x \geqslant 0 \\ -\beta m(|x|) & x < 0 \end{cases} \tag{2}$$

The $\beta \in (0,1]$ parameter specifies the degree of compression.

The third assumption is that number magnitudes are compared by first computing the mathematical distance between them.

$$d(x,y) = s(x) - s(y) \tag{3}$$

The fourth assumption is that comparison time is inversely proportional to this distance.

$$RT(x,y) \propto -\log(d(x,y)) \tag{4}$$

These equations trace back to classic theories of number comparison (Moyer & Landauer, 1967; Shepard et al., 1975), and ultimately to classic accounts of decision making (Welford, 1960).

A.2. Empirical evaluation and sensitivity analysis

The reflection model was empirically evaluated by computing the correlation between model and human comparison times. Grid search was used to estimate the values of α and β that maximized this correlation. Table A1 shows the results for Experiments 1 and 2. In both cases, the model provides a good quantitative account of the data.

The model contains two free parameters that are estimated from the data: the curvature of mental number lines (α) and the compression of the negative number line (β). Separate sensitivity analyses for Experiments 1 and 2 revealed that the model could be reduced to a single free parameter (Table A1). If β is set to a default value of 1 and grid search is used to estimate α, then the reduced model fits the data poorly. However, if α is set to a default value of 1 and grid search used to estimate β, the fit remains very good. Thus, the reflection model only requires estimation of a single parameter β specifying the compression of the negative number line.

A.3. Judgment effects

The reflection model specifies a core representation for integers. It does not make direct predictions about judgment effects, which depend on the association between the core representation and either a cognitive or physical operation. However, it can make indirect predictions when augmented with auxiliary assumptions, for example that the natural number line is associated with the greater predicate and the negative number line with the lesser predicate. Under this assumption, the model predicts the general pattern of semantic congruence effects observed in Experiments 1 and 2: a standard semantic congruence effect for positive and mixed-positive comparisons and an inverse semantic congruence effect for negative and mixed-negative comparisons. Note that the augmented reflection model fails to account for the semantic congruence effect for zero comparisons: It predicts a standard semantic congruence effect for zero-positive comparisons (because positive numbers are associated with the greater predicate) and an inverse semantic congruence effect for zero-negative comparisons (because negative numbers are associated with the lesser predicate), neither of which was observed in Experiments 1 and 2. This is another instance of the general finding that zero behaves differently than other numbers.

Appendix B. The symbolic-magnitude model

This appendix formalizes the symbolic-magnitude model, evaluates its quantitative fit to the adult data, reports a sensitivity analysis demonstrating the sufficiency of a reduced model with only one parameter, and summarizes its qualitative account of judgment effects.

B.1. Model formalization

The symbolic magnitude of the numbers $\{-N,\ldots,0,\ldots,N\}$ can be formalized as follows.

Number representations consist of $N+1$ features. The first assumption specifies how psychological magnitudes m_i are encoded.

$$m_i = i^{-\alpha} \quad i \geqslant 0 \tag{5}$$

As in the reflection model, the $\alpha \in (0,1]$ parameter specifies the scaling of psychological magnitudes. The second assumption specifies how signed integer magnitudes are constructed from psychological magnitudes. The sign feature s_0, the symbolic component of the representation, is defined as:

$$s_0(x) = \begin{cases} 0 & x \geqslant 0 \\ 1 & x < 0 \end{cases} \tag{6.1}$$

The magnitude features s_1 through s_N are defined as:

$$s_i(x) = \begin{cases} m_i & 0 \leqslant i \leqslant x \\ \beta m_i & -i < x < 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.2}$$

As in the reflection model, the $\beta \in (0,1]$ parameter specifies the compression of negative magnitudes relative to positive magnitudes.

Number representations are compared using similarity-based processing (Nosofsky, 1984; Shepard, 1987). The fea-

Table A1
Fit of the reflection model to the data of Experiments 1 and 2.

	Full model	Reduced models[a]	
		α Free, $\beta = 1$	$\alpha = 1$, β Free
Experiment 1			
Fit to data	$r(33) = 0.94$, $p < .001$	$r(33) = 0.63$, $p < .001$	$r(33) = 0.92$, $p < .001$
Comparison to full model		$p < .001$	*ns*
Experiment 2			
Fit to data	$r(92) = 0.91$, $p < .001$	$r(92) = 0.61$, $p < .001$	$r(92) = 0.90$, $p < .001$
Comparison to full model		$p < .001$	*ns*

[a] Parameter α reflects the scaling of number lines and parameter β the compression of the negative number line relative to the natural number line.

Table B1
Fit of the symbolic-magnitude model to the data of Experiments 1 and 2.

	Full model	Reduced models[a]	
		α Free, $\beta = 1$	$\alpha = 1$, β Free
Experiment 1			
Fit to data	$r(33) = 0.96$, $p < .001$	$r(33) = 0.63$, $p < .001$	$r(33) = 0.96$, $p < .001$
Comparison to full model		$p < .001$	*ns*
Experiment 2			
Fit to data	$r(92) = 0.92$, $p < .001$	$r(92) = 0.61$, $p < .001$	$r(92) = 0.92$, $p < .001$
Comparison to full model		$p < .001$	*ns*

[a] Parameter α reflects the scaling of magnitude feature values and parameter β the compression of negative magnitudes relative to positive magnitudes.

tural difference between two number representations is defined using a city-block metric.

$$d(x,y) = \sum_{j=0}^{N} |s_j(x) - s_j(y)| \quad (7)$$

Response time is proportional to the similarity of two number representations, defined as a negative exponential function of their featural difference.

$$RT(x,y) \propto e^{-d(x,y)} \quad (8)$$

B.2. Empirical evaluation and sensitivity analysis

The symbolic-magnitude model provides a good quantitative account of Experiments 1 and 2, as summarized in Table B1. Sensitivity analyses indicate that the model's fit depends critically on the compression of negative magnitudes (β), but is relatively independent of the scaling of psychological magnitudes (α).

B.3. Judgment effects

The symbolic-magnitude model can make indirect predictions about the semantic congruence effect under the auxiliary assumption that a value of 0 for the sign feature s_0 is associated with the greater predicate and a value of 1 with the lesser predicate. Under this assumption, it correctly predicts the general results of Experiments 1 and 2: a standard semantic congruence effect for positive and mixed-positive comparisons, and an inverse semantic congruence effect for negative and mixed-negative comparisons. However, it incorrectly predicts standard and inverse effects for zero-positive and zero-negative comparisons, respectively. This is another instance of the general finding that zero behaves differently.

References

Anderson, J. R. (1978). Arguments concerning representations for mental imagery. *Psychological Review, 85*, 249–277.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369–406.

Anderson, J. R., Corbett, A. T., Koedinger, K., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of Learning Sciences, 4*, 167–207.

Ansari, D., Garcia, N., Lucas, E., Hamon, K., & Dhital, B. (2005). Neural correlates of symbolic number processing in children and adults. *NeuroReport, 16*, 1769–1773.

Banks, W. P., Fujii, M., & Kayra-Stuart, F. (1976). Semantic congruity effects in comparative judgments of magnitudes of digits. *Journal of Experimental Psychology: Human Perception and Performance, 2*, 435–447.

Banks, W. P., & Hill, D. K. (1974). The apparent magnitude of number scaled by random production. *Journal of Experimental Psychology Monograph, 102*, 353–376.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577–609.

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*, 434–452.

Crowley, M. L., & Dunn, K. A. (1985). On multiplying negative numbers. *Mathematics Teacher, 78*, 252–256.

Curtis, D. W., & Fox, B. E. (1969). Direct quantitative judgments of sums and a two-stage model for psychophysical judgments. *Perception & Psychophysics, 5*, 89–93.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*, 371–396.

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition, 43*, 1–29.

Duncan, E. M., & McFarland, C. E. Jr., (1980). Isolating the effects of symbolic distance and semantic congruity in comparative judgments: An additive-factors analysis. *Memory & Cognition, 8*, 612–622.

Euler, L. (1755/2000). *Foundations of differential calculus. Translated by J.D. Blanton*. New York: Springer-Verlag.

Fischer, M. H. (2003). Cognitive representation of negative numbers. *Psychological Science, 14*, 278–282.

Fischer, M. H., & Rottman, J. (2005). Do negative numbers have a place on the mental number line? *Psychology Science, 47*, 22–32.

Franklin, M. S., Jonides, J., & Smith, E. E. (2009). Processing of order information for numbers and months. *Memory & Cognition, 37*, 644–654.

Ganor-Stern, D., & Tzelgov, J. (2008). Negative numbers are generated in the mind. *Experimental Psychology, 55*, 157–163.

Gevers, W., Santens, S., Dhooge, E., Chen, Q., Van den Bossche, L., Fias, W., et al. (2010). Verbal-spatial and visuospatial coding of number–space interactions. *Journal of Experimental Psychology: General, 139*, 180–190.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences, 2*, 1–55.

Harnad, S. (1990). The symbol grounding problem. *Physica D, 42*, 335–346.

Hativa, N., & Cohen, D. (1995). Self learning of negative number concepts by lower division elementary students through solving computer-provided numerical problems. *Educational Studies in Mathematics, 28*, 401–431.

Hayes, B., & Stacey, K. (1999, July). Teaching negative number using integer tiles. In *Proceedings of the 22nd annual conference of the mathematics education research group of Australasia (MERGA), Adelaide, Australia.*

Holyoak, K. J. (1978). Comparative judgments with numerical reference points. *Cognitive Psychology, 10*, 203–243.

Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reasoning*. Chicago, IL: University of Chicago Press.

Just, M. A., & Varma, S. (2007). The organization of thinking: What functional brain imaging reveals about the neuroarchitecture of cognition. *Cognitive, Affective, and Behavioral Neuroscience, 7*, 153–191.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science, 320*, 454–455.

Klein, F. (1925). *Elementary mathematics from an advanced standpoint: Part 1. Arithmetic-algebra-analysis*. Translated by E. R. Hedrick and C. A. Noble from the Springer 3rd (1925) edition, New York: Dover.

Knops, A., Thirion, B., Hubbard, E. M., Michel, V., & Dehaene, S. (2009). Recruitment of an area involved in eye movements during mental arithmetic. *Science, 324*, 1583–1585.

Kosslyn, S. M. (1981). The medium and the message in mental imagery: A theory. *Psychological Review, 88*, 46–66.

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.

Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. NY, NY: Basic Books.

Landy, D., & Goldstone, R. L. (2007). How abstract is symbolic thought? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 33*, 720–733.

Lewis, M. R., & Varma, S. (2010). *Learning environments, numerical representations, and the emergence of distance and magnitude effects in connectionist models of digital comparison*. Poster presented at the 32nd annual meeting of the cognitive science society, Portland, OR.

Liebeck, P. (1990). Scores and forfeits, and intuitive model for integers. *Educational Studies in Mathematics, 21*, 221–239.

Luria, A. R. (1966). *Higher cortical functions in man*. London: Tavistock.

Marr, D. (1982). *Vision: A computational investigation into the human representation and presentation of visual information*. New York: W.H. Freeman.

Mesulam, M.-M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language and memory. *Annals of Neurology, 28*, 597–613.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167–202.

Miller, G., & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge, MA: Harvard University Press.

Montessori, M. (1966). *The discovery of the child*. Kalakshetra Publications.

Moreno, R., & Mayer, R. E. (1999). Multimedia-supported metaphors for meaning making in mathematics. *Cognition and Instruction, 17*, 215–248.

Moyer, R. S., & Landauer, T. K. (1967). The time required for judgments of numerical inequality. *Nature, 215*, 1519–1520.
Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
Newman, J. (1956). *The world of mathematics*. New York: Simon and Schuster.
Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 10*, 104–114.
Nuerk, H.-C., Iversen, W., & Willmes, K. (2004). Notational modulation of the SNARC and MARC (linguistic markedness of response codes) effect. *The Quarterly Journal of Experimental Psychology, 57A*, 835–863.
Nuerk, H.-C., Weger, U., & Willmes, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition, 82*, B25–B33.
Parkman, J. M. (1971). Temporal aspects of digit and letter inequality judgments. *Journal of Experimental Psychology, 91*, 191–205.
Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron, 44*, 547–555.
Pinel, P., Dehaene, S., Rivière, D., & Le Bihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *NeuroImage, 14*, 1013–1026.
Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron, 41*, 983–993.
Pinhas, M., & Fischer, M. H. (2008). Mental movements without magnitude? A study of spatial biases in symbolic arithmetic. *Cognition, 109*, 408–415.
Prather, R. W., & Alibali, M. W. (2008). Understand and using principles of arithmetic: Operations involving negative numbers. *Cognitive Science, 32*, 445–457.
Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin, 132*, 416–442.
Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review, 88*, 16–45.
Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology, 83*, 274–278.
Rivera, S. M., Reiss, A. L., Eckert, M. A., & Menon, V. (2005). Developmental changes in mental arithmetic: Evidence for increased specialization in the left inferior parietal cortex. *Cerebral Cortex, 15*, 1779–1790.
Santens, S., & Gevers, W. (2008). The SNARC effect does not imply a mental number line. *Cognition, 108*, 263–270.
Sasaki, Y., Vanduffel, W., Knutsen, T., Tyler, C., & Tootell, R. (2005). Symmetry activates extrastriate visual cortex in human and nonhuman primates. *Proceedings of the National Academy of Sciences United States of America, 102*, 3159–3163.
Schwartz, D. L., Blair, K., P., & Tsang, J. M. (in press). How to build an educational neuroscience: Two approaches with concrete instances. *British Journal of Educational Psychology*.
Schwarz, B. B., Kohn, A. S., & Resnick, L. B. (1993). Positives about negatives: A case study of an intermediate model for signed numbers. *Journal of the Learning Sciences, 3*, 37.
Sekuler, R., & Mierkiewicz, D. (1977). Children's judgments of numerical inequality. *Child Development, 48*, 630–633.
Shaki, S., & Petrusic, W. M. (2005). On the mental representation of negative numbers: Context-dependent SNARC effects with comparative judgments. *Psychonomic Bulletin & Review, 12*, 931–937.
Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science, 237*, 1317–1323.
Shepard, R. N., Kilpatric, D. W., & Cunningham, J. P. (1975). The internal representation of numbers. *Cognitive Psychology, 7*, 82–138.
Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701–703.
Stevens, S. S. (1970). Neural events and the psychophysical law. *Science, 170*, 1043–1050.
Thompson, P. W., & Dreyfus, T. (1988). Integers as transformations. *Journal for Research in Mathematics Education, 19*, 115–133.
Tsang, J. M., Rosenberg-Lee, M., Blair, K. P., Schwartz, D. L., & Menon, V. (2010, June). *Near symmetry in a number bisection task yields faster responses and greater occipital activity*. Poster presented at the 16th annual meeting of the organization for human brain mapping, Barcelona, Spain.
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.
Varela, F. J., Thompson, E., & Rosch, E. (1992). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.
Varma, S., Harris, S., Schwartz, D. L., & Martin, T. (2009, April). *The development of the integer concept: A cross-sectional study*. Paper presented at the 2009 annual meeting of the American educational research association, San Diego, CA.
Varma, S., Schwartz, D. L., Lindgren, R., & Go, J. (2007). Integer comparison and the inverse distance effect. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th annual cognitive science society* (pp. 1875). Austin, TX: Cognitive Science Society.
Verguts, T., & De Moor, W. (2005). Two-digit comparison: Decomposed, holistic, or hybrid? *Experimental Psychology, 52*, 195–200.
Welford, A. T. (1960). The measurement of sensory-motor performance: Survey and reappraisal of twelve years progress. *Ergonomics, 3*, 189–230.
Zorzi, M., & Butterworth, B. (1999). A computational model of number comparison. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the 21st annual conference of the cognitive science society* (pp. 772–777). Mahwah, NJ: Erlbaum.