Possible number systems

Lance J. Rips · Samantha Thompson

Abstract Number systems—such as the natural numbers, integers, rationals, reals, or complex numbers—play a foundational role in mathematics, but these systems can present difficulties for students. In the studies reported here, we probed the boundaries of people’s concept of a number system by asking them whether “number lines” of varying shapes qualify as possible number systems. In Experiment 1, participants rated each of a set of number lines as a possible number system, where the number lines differed in their structures (a single straight line, a step-shaped line, a double line, or two branching structures) and in their boundedness (unbounded, bounded below, bounded above, bounded above and below, or circular). Participants also rated each of a group of mathematical properties (e.g., associativity) for its importance to number systems. Relational properties, such as associativity, predicted whether participants believed that particular forms were number systems, as did the forms’ ability to support arithmetic operations, such as addition. In Experiment 2, we asked participants to produce properties that were important for number systems. Relational, operation, and use-based properties from this set again predicted ratings of whether the number lines were possible number systems. In Experiment 3, we found similar results when the number lines indicated the positions of the individual numbers. The results suggest that people believe that number systems should be well-behaved with respect to basic arithmetic operations, and that they reject systems for which these operations produce ambiguous answers. People care much less about whether the systems have particular numbers (e.g., 0) or sets of numbers (e.g., the positives).

Keywords Number systems · Mathematical cognition · Number representation

A tribute and an introduction

Ed Smith was a person of great warmth and intellectual clarity. A true generalist, he published in many areas of cognitive psychology, as well as in cognitive neuroscience, social psychology, and even (in a few early articles) clinical psychology. He brought to these projects a clear-headed and unpretentious viewpoint, evident in both his conversation and his articles. Among his many other accomplishments, Ed was a pioneer in research on concepts in cognitive psychology. Along with Eleanor Rosch, he was the first to document differences in typicality among types of natural categories: Some birds (e.g., sparrows) are more typical than other birds (e.g., emus), and some furniture (e.g., sofas) is more typical than other furniture (e.g., picture frames). E. E. Smith and Medin’s (1981) book, Categories and Concepts, provides an important statement of these findings. Although many rival models now exist for these effects, typicality gradients continue to provide foundational results for theories in this area.

We now know that typicality effects are everywhere. Even well-defined categories show typicality differences. For example, some odd numbers (e.g., 3) are judged as being more typical than other odd numbers (e.g., 4,284,647), as Armstrong, Gleitman, and Gleitman (1983) first demonstrated. Odd numbers are well-defined concepts, at least in the sense that both 3 and 4,284,647 are definitely odd, not vaguely so. Thus, the fact that people believe that 3 is a more typical odd number than 4,284,647 suggests that typicality gradients are not diagnostic of a category’s lack of precise definition. A theory of human concepts should account for both their well-defined or ill-defined structure and their typicality differences (see Osherson & Smith, 1981, for one attempt to do so).
However, beyond the dialectical role that mathematical concepts have played in debates about typicality, they have not been the topic of much research in the concepts-and-categories area. Developmentalists, of course, have devoted attention to how children learn numerical concepts, such as the positive integers, because of the importance of these concepts in school mathematics (see Rips, Bloomfield, & Asmuth, 2008, for a review). But the research on adult math concepts has focused mainly on how people perform simple arithmetic, such as addition or number comparison. Relatively few studies have looked at the ways that adults cope with novel mathematical systems or notations.

In this article, we begin to redress this lack of research by exploring people’s concepts of number systems. We presented participants with number lines—varying in how the numbers were bounded and how they could be compared—and asked the participants for their judgments of whether the number lines could represent possible number systems. We then tried to predict these judgments from the mathematical properties that the lines supported. Probing these intuitions provided clues as to people’s assumptions about the nature of numbers and of mathematics.

**Number systems: Conceptions and preconceptions**

Number systems—for example, the natural numbers, integers, rationals, reals, and complex numbers—provide frameworks for mathematics. They define what qualifies as a number within their domains, support arithmetic and other mathematical operations, and provide the groundwork for higher mathematics (Feferman, 1964). Because number systems delineate correct inferences about numbers, people’s mathematical reasoning depends for its accuracy on whether it conforms to the systems’ underlying principles. For instance, whether you can always subtract a larger number from a smaller depends on whether you are in the realm of the natural numbers or the integers. Whether you can always divide one whole number by another depends on whether you are in the realm of the integers or the rationals.

For these reasons, large portions of elementary and secondary education in mathematics are taken up teaching the properties of individual number systems, beginning with the natural numbers and graduating to the more complicated ones. The National Council of Teachers of Mathematics (2000) lists among its expectations for grades 9–12 that students “compare and contrast the properties of numbers and number systems, including the rational and real numbers, and understand complex numbers as solutions to quadratic equations that do not have real solutions” and “understand vectors and matrices as systems that have some of the properties of the real-number system.”

These expectations are not always easy for students to meet. Research suggests that students in early grades attempt to understand fractions by unreliably adapting strategies suited to their previous knowledge of the positive integers (Behr, Wachsmuth, Post, & Lesh, 1984; Hartnett & Gelman, 1998; C. L. Smith, Solomon, & Carey, 2005). Even adults make systematic errors when faced with novel mathematical systems or notations, sometimes carrying over their knowledge of typical arithmetic procedures to systems in which they no longer apply (Ben-Zeev, 1995). These difficulties in learning mathematics indicate that students’ preconceptions about number systems can interfere with their progress. To help students around these barriers, teachers may find it useful to know the students’ intuitive beliefs about the properties and relations that numbers have to possess, and the properties that disqualify something as a number. For example, we will suggest later, on the basis of the results of the present experiments, that college student may overvalue the ability to compare any two numbers (as being less than, equal to, or greater than one another) within a number system. Teaching them that this property is not a necessary one for number systems may make it easier for them to learn systems like the complex numbers, in which that property does not hold.

In addition to these practical implications, exploring people’s concepts of number systems can provide clues about the mental representation of number. Because typical number systems contain an infinite set of elements, people cannot represent all numbers in such systems as individual units in memory, but must derive them from underlying principles. Although mathematical axioms are a formal statement of these principles, mathematicians usually devise these axioms with a view toward their convenience in proofs rather than their psychological plausibility. People’s intuitions about number systems may shed light on the mental structure of this knowledge.

**Envisioning number systems**

In the experiments reported here, we used a set of novel number “lines” to elicit intuitions about number systems. The research followed Shepard, Kilpatrick, and Cunningham’s (1975) pioneering study, which used visual symbols for mathematical objects to probe people’s mental representations of numbers. However, instead of symbols for individual numbers, we used diagrams of sets of numbers—number lines—since our interest was in entire number systems. Recent studies have used number lines to explore how European, US, and Amazonian natives represent numbers (for Amazonians, see Dehaene, Izard, Spelke, & Pica, 2008; for European and US children, see, e.g., Barth & Paladino, 2011; Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010; Booth & Siegler, 2006; Ebersbach, Luwel, Frick, Ongena, & Verschaffel, 2008; Geary, Hoard, Nugent, & Byrd-Craven, 2008; Moeller, Pixner, Kaufmann, & Nuerk, 2009; and Siegler & Opfer, 2003;

The present study focuses on adults’ ideas about the structure of number systems, rather than on the meaning of individual numerals. For this reason, participants saw only the shape of the number line, either unlabeled (Exp. 1), with a label for 0 alone (Exp. 2), or with tick marks for other, unidentified numbers (Exp. 3). Participants rated the number lines according to whether they were potential number systems, and these ratings comprised the dependent variable for the experiments. Figure 1 previews the lines that we used in Experiment 2. The lines differed in their segmentation and boundaries, and these factors were independent variables, as we are about to discuss. In addition, we compared types of mathematical properties as predictors of the number-system ratings: whether the number lines were consistent with math operations (e.g., addition), relations (e.g., commutativity), or elements (e.g., a first number).

**Comparability and boundedness as properties of number systems**

The lines in Fig. 1 vary along two dimensions that we hypothesized might influence people’s ideas about the acceptability of sets of numbers as number systems. The first (vertical) dimension is the extent to which individual numbers are *comparable*: that is, whether all numbers or only subsets of numbers can be compared as being less than, equal to, or greater than one another. The second (horizontal) dimension is the way in which the numbers are *bounded*: that is, whether the numbers have a starting or stopping point or are, instead, infinite in a positive or negative direction.

The stimulus diagrams in Fig. 1 include items with traditional shapes (single lines) and nontraditional shapes (double lines, step-shaped lines, and branching lines). These types of shapes appear in separate rows of the figure. Because these diagrams are not line segments in the geometric sense, we will refer to them as *number forms* (following Galton, 1880, 1881) rather than *number lines*. To denote the individual number forms, we will use the row and column labels in the figure. For example, the form in the upper-left corner is the linear unbounded item, and the form in the lower-right corner is the backward-branching circular item. These diagrams appear to be the simplest that could plausibly illustrate number systems and that vary in mathematically important ways.

The rows of the figure depict forms that vary in their internal comparability, according to the relations less than (<), greater than (>), or equal to (=). The linear and step forms allow comparisons of any two numbers: For any numbers \(x\) and \(y\) on the forms, either \(x < y\), \(x > y\), or \(x = y\). The remaining
forms allow comparisons for only subsets of numbers. The bilinear forms prohibit comparisons between the two lines, and the branching forms prohibit comparisons between branches. (Participants learned that a number on the upper line or branch was neither less than, greater than, or equal to one on the lower line or branch.) The linear and step systems, in which any two numbers can be compared, make for more straightforward reasoning and computation. Most of our college-student participants had encountered complex numbers in their earlier math training, and in this system some number pairs (e.g., 1 and i) are not comparable. Nevertheless, students may find such number systems to be less intuitive than ones that allow all possible comparisons, as we mentioned earlier.

The columns of Fig. 1 represent ways in which sets of numbers can be bounded. Columns with forms that run into the right side of the surrounding frame (columns 1 and 2) continue to infinity in the positive direction, and those with forms that run into the left side (columns 1 and 3) continue to infinity in the negative direction. (The instructions explained these conventions to the participants.) Thus, the forms in the first column are unbounded, those in the second column are bounded below (but not above), and so on. Sets that are bounded below may seem more natural to participants than ones that are bounded above. For example, the positive integers may seem a more reasonable number system than the negative integers, despite the isomorphism between them. (The circular forms in the last column raise issues of comparison as well as of boundedness, and we will keep this fact in mind in interpreting the results.) Comparability and boundedness do not exhaust the dimensions that are important in evaluating sets of numbers as number systems, as we are about to see; however, they do have important implications for reasoning about numbers.

Other properties of number systems and number forms

To get a more detailed look at people’s ideas about numbers, we also asked participants to evaluate a set of mathematical properties for their importance to number systems. We then used these property ratings to predict their assessments of the number forms. The properties in Experiments 1 and 3 came from mathematical sources. Those in Experiment 2 came from a preliminary study in which we asked a new group of participants to play the role of anthropologists studying the numbers of a native people. The participants were to produce properties that they believed would be useful to know in deciding whether the natives’ numbers formed a number system. We used the most frequently mentioned properties from this group in place of the textbook properties as predictors of the ratings of the number forms.

The properties that we used in the experiments—both the textbook properties and the participant-generated properties—varied in whether they emphasized specific numbers (and sets of numbers) or interrelations among the numbers. Among the former were characteristics such as whether there was a first number or whether there was an infinite set of numbers. Among the latter were properties that highlighted arithmetic operations (e.g., whether addition was defined for the numbers) and relations (e.g., associativity and commutativity). According to one current theory in the philosophy of mathematics, the existence of numbers depends on their position within a relational structure (e.g., Resnik, 1997; Shapiro, 1997). For example, the natural number 3 is defined by the position that it occupies in the natural number sequence between 2 and 4. From this point of view, we predicted that properties mentioning relations and operations would do a better job than properties naming specific elements in predicting the ratings of the number forms as possible number systems.

The forms in Fig. 1 differ not only in their mathematical properties but also in their perceptual ones. For example, the backward-branching circular form is more complex perceptually than the linear bounded form. Because these perceptual factors could influence participants’ possible-number-system judgments, we asked a separate group of participants to rate the forms’ complexity and their similarity to a straight-line segment. We then could use these ratings to control for perceptual factors in analyzing the more mathematical variables.

An overview of the experiments

The main goal of these experiments was to determine what factors affect people’s ideas of number systems. All of the experiments varied comparability and boundedness, but they differed in the additional information that the number forms contained. In Experiment 1, we used lines with no numerical labels or tick marks. In Experiment 2, we added a label and a tick mark for zero (as in Fig. 1). Experiment 3 included tick marks for all numbers (see Fig. 4). To help interpret these judgments at a finer level, participants also read a list of mathematical properties, such as commutativity and associativity, and rated each property’s importance for number systems. We used these ratings to determine whether participants’ concepts of number systems depend on operations on numbers, relations between numbers, or the presence of individual numbers or sets.

The participants were college students already familiar with systems like the integers and rational numbers. Thus, their intuitions about the nature of number systems might be shaped both by their familiarity with these particular systems and by their more abstract understanding of number. Separating these sources of knowledge is difficult, because knowledge of specific examples can mimic abstract knowledge, as research on exemplar models in categorization has shown (see, e.g., E. E. Smith & Medin, 1981). The use of novel number forms, however, provided some assurance that the participants were
extending their knowledge to cases that they had not previously encountered, rather than simply giving rote answers to familiar problems.

Experiment 1: Number forms as possible number systems

Participants in one part of this experiment viewed a series of number forms one at a time, and they rated these forms as possible number systems. Instructions at the beginning of the session reminded participants, “You are already acquainted with number systems such as the integers, rational numbers (fractions), and real numbers. We are interested in your ideas about what makes something a number system and what the characteristics of number systems are.” The number forms were similar to those of Fig. 1; however, they did not show the position of zero (no tick mark or label for zero appeared). To clarify the meanings of the forms, the instructions stated that the diagrams were similar to number lines in that they depicted sets of numbers increasing in size in the direction of the arrows. The instructions also stated that some of the diagrams showed two different number lines that were not connected, and that numbers on one line were not smaller, larger, or equal to the numbers on the other. Similarly, in the branching structures, the numbers on each branch could be compared to the numbers on the shared trunk, but not to the numbers on the other branch. Participants also learned about the conventions governing boundedness, described earlier.

The experimental session included a second part, in which participants rated each of a group of mathematical properties for their importance to number systems. Table 1 contains a list of these properties as they appeared in the experiment. (The labels in Table 1, such as Associativity, and the headings, such as Operations, are for convenience in referring to the properties; the participants did not see them.) We selected the properties as being ones that are true of some (and, in some cases, possibly all) number systems, and they fall in three categories in the table. The relational properties, such as commutativity and associativity, hold among pairs or triples of numbers; the element properties, such as the presence of a first element or of an infinite set of elements, posit the existence of particular numbers or sets; and the operation properties, such as addition or division, specify the arithmetic operations that the system supports. The instructions stated, “Some properties are so important that all number systems must have them, while other properties may not be important to any number system.” Participants rated the importance of each property to number systems and later indicated which properties were true of which number forms. The purpose of this part of the experiment was to determine which types of properties predicted the possible-number-system ratings.

Table 1 Stimulus properties in Experiment 1, with mean ratings of importance to number systems and correlations of the weighted property ratings to the ratings of number forms as possible number systems

<table>
<thead>
<tr>
<th>Relations</th>
<th>Mean Rated Importance</th>
<th>Correlation With Number-System Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Closure</td>
<td>4.34</td>
<td>.59**</td>
</tr>
<tr>
<td>Trichotomy</td>
<td>7.36</td>
<td>.78**</td>
</tr>
<tr>
<td>Inequality preserved</td>
<td>4.28</td>
<td>.70**</td>
</tr>
<tr>
<td>First element</td>
<td>5.19</td>
<td>.25</td>
</tr>
<tr>
<td>Last element</td>
<td>3.30</td>
<td>.09</td>
</tr>
<tr>
<td>Infinity</td>
<td>5.69</td>
<td>.38</td>
</tr>
<tr>
<td>Successor</td>
<td>5.68</td>
<td>.01</td>
</tr>
<tr>
<td>Density</td>
<td>5.36</td>
<td>.90**</td>
</tr>
<tr>
<td>Operations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Addition</td>
<td>6.47</td>
<td>.67**</td>
</tr>
<tr>
<td>Multiplication</td>
<td>6.00</td>
<td>.65**</td>
</tr>
<tr>
<td>Subtraction</td>
<td>6.36</td>
<td>.73**</td>
</tr>
<tr>
<td>Division</td>
<td>5.28</td>
<td>.90**</td>
</tr>
</tbody>
</table>

*p < .05, ** p < .01, df = 23
A separate set of participants provided ratings of the number forms’ perceptual properties. These participants viewed the same number forms, but without knowing that the forms represented numbers. They rated the forms’ complexity and similarity to a straight line, and we used these ratings in the analysis to control these perceptual factors.

Method

Two groups of participants took part in this experiment. One group rated the number forms as being possible number systems and rated the properties in Table 1 for their importance to number systems. A second group rated the number forms for their perceptual complexity and perceptual similarity to a simple line segment.

Number form and property ratings A computer screen displayed the 25 number forms in a new random order for each participant. For each form, participants indicated their rating by clicking a numeral on a vertical 1–8 scale, with 1 labeled NOT a number system and 8 labeled Number system. The instructions stated, “If you think the diagram illustrates a clear example of a number system, give the diagram a rating of 8 by clicking on the scale at the ‘8’ position. If you think the diagram could not possibly illustrate a number system, click on the ‘1’ position. Use intermediate numbers for intermediate possibilities.”

For the property ratings, participants saw each of the 15 properties from Table 1, again presented one at a time in random order on the computer screen. We asked participants to consider number systems that they already knew, such as the integers, rationals, and real numbers, as well as possible number systems that mathematicians had not yet invented. They then rated the importance of each property in relation to all possible number systems. The wording of each property was exactly as it appears in Table 1 (but without the labels). Participants again responded by clicking on a 1-to-8 scale, but this time the scale label for 1 read need not be part of any number system, and for 8, must be part of every number system. As a guide to the meaning of the properties, the instructions told participants that S stood for the set of numbers in the candidate number system, and x, y, and z for any numbers in S. The instructions also informed participants that “+” stood for addition and “*” for multiplication in the usual way, but added that “sometimes we will also need to talk about an arbitrary operation on numbers, which could be addition, multiplication, subtraction, division, or any other operation. We will use the symbol # for such an operation.”

Following the property ratings, the participants identified the number forms for which each of the Table 1 properties was true. The participants received a booklet, on each page of which one of the properties appeared at the top. Below the property was an array of the number forms, similar to the array in Fig. 1. The participants indicated whether the property held for each form by circling “yes” or “no” beneath the form. The booklet pages were randomly ordered, with a new order for each participant. Linking the properties and forms in this way allows us to use the properties’ importance to predict the possible-number-system ratings, as we describe in the Results and Discussion section. The order of the tasks was balanced: Half of the participants rated the number forms, rated the properties’ importance, and then indicated the connections between the forms and the properties; and the remaining participants rated the properties, indicated the property–form connections, and finally rated the number forms. At the end of the session, participants recorded the math classes they had taken in high school and college, and they commented on the methods that they had followed in rating the number forms. Participants took from 30 to 50 min to complete the experiment.

Similarity and complexity ratings Participants’ judgments about whether a number form is a possible number system could be due to the form’s perceptual complexity or to the perceptual similarity of the form to a standard straight line. For this reason, an independent group of participants rated the perceptual complexity of each number form and the similarity of each form to a single line.

For the complexity ratings, the instructions informed participants that they would be seeing a series of diagrams that varied in their perceptual complexity, but the instructions made no mention of the fact that the lines might represent numbers. A computer presented the 25 forms one at a time in the same way as in the possible-number-system task. Participants recorded their complexity ratings by clicking a numeral on a 1-to-8 point scale, with 1 marked very simple perceptually and 8 marked very complex perceptually. For the similarity ratings, the instructions mentioned that the participants would be seeing diagrams that varied in their similarity to straight lines, but again the instructions did not refer to numbers. The instructions also showed participants a sample diagram that resembled the unbounded linear form in Fig. 1. The computer again presented the 25 forms, and the participants rated the similarity of each form to the sample on a scale from 1 (very dissimilar) to 8 (very similar). The session lasted 10–15 min.

Participants The participants in the main experiment were 36 students from an introductory psychology class at Northwestern University. They received credit toward a course requirement in exchange for performing these tasks. All participants had taken a calculus course in either high school or college. Ten had also taken one or more advanced courses, usually multivariate
calculus or linear algebra. None reported courses in abstract algebra, set theory, or number theory.

An additional 32 students participated in the similarity and complexity ratings. These students were from the same population as those in the main experiment, but none had participated in that study. Of these participants, 17 completed the similarity ratings before the complexity ratings, and 15 completed these tasks in the reverse order.

Results and discussion

In examining the results from this experiment, we will first look at the mean ratings of the number forms as possible number systems. We can then see to what extent we can predict these ratings on the basis of their mathematical properties and of their complexity and similarity to a straight line. A preliminary analysis of variance (ANOVA) showed no reliable effect of the order in which participants completed the rating tasks, and the results reported here are collapsed over the two orders.

**Ratings of the number forms as possible number systems** Mean ratings of the number forms showed clear effects of their internal comparability (the row variable in Fig. 1). The effects of boundedness were small, except in the case of the circular structures in the last column of Fig. 1. Figure 2 plots these ratings for all of the number forms, and the marginal means for the comparability factor appear in the first data column of Table 2. Participants gave linear structures the highest mean rating, as might be expected from the common textbook use of line segments for number lines. The bilinear and step forms, however, also received relatively high ratings. Although our participants had probably never considered two entirely unconnected sets of numbers as composing a potential number system, they apparently did not see the bilinear form as being completely disqualified from serving this function. The step forms consisted of a single line, and participants might have treated them in the same way as the linear forms. However, the step forms’ vertical dimension probably caused uncertainty. The instructions told participants that “the direction of the arrows indicates the ordering of the numbers” but gave them no further information about the vertical differences. The contrast between the step and linear forms may have hinted that the steps had some additional import. We will see further evidence for this possibility in Experiments 2 and 3.

The branching forms received relatively low, and nearly identical, ratings. Participants took branching as a more serious defect in potential number systems than a complete split among the numbers (as in the bilinear forms). These participants may have worried that the branching forms could create conflicts in the results of arithmetic operations. For example, if two numbers along the main “trunk” of the forward-branching form added to a number beyond the trunk, the result would seem to have two different and incomparable values. In line with the differences among the means, an ANOVA produced a significant effect of comparability, $F(4, 140) = 15.95$, $MSE = 9.04$, $p < .001$. A multiple-range test (REGW with $\alpha = .05$; Kirk, 1995) showed that the mean for the linear forms was significantly higher than those for the bilinear or step forms, and that the means for the bilinear and step forms were in turn significantly higher than those for the branching forms. No further pairs of means differed significantly. Thus, the connectedness of the branching forms appears to be less important than the potential ambiguity that they create for operations such as addition and subtraction.

Figure 2 shows that similar ratings were associated with upper and lower bounding. The means were 5.73 for the unbounded forms, 5.67 for forms bounded above, 5.49 for forms bounded below, and 5.47 for forms bounded both above and below. Only the circular forms had noticeably lower ratings: 3.53 on the 1-to-8 scale. The ANOVA produced a significant effect of boundedness [$F(4, 140) = 22.61$, $MSE = 6.84$, $p < .001$], but a multiple-range test showed a difference only between the circular forms and the remaining ones.

The circular forms provide a special case: Ratings were uniformly low for these items and did not depend on other aspects of their shape. This produced a significant interaction between comparability and boundedness, which is evident in Fig. 2, $F(16, 560) = 3.75$, $MSE = 1.58$, $p < .001$. This interaction may reflect the fact that the circular forms not only bound the numbers in their domains but also complicate comparisons: The circular structures imply that a number $x$ can be seen as being either less than or greater than a different number $y$. Convention could impose an ordering on such forms, in the same way that 1 a.m. is conventionally earlier than 5 a.m., despite the cyclical structure of clock times. However, participants in this experiment had no such convention at hand. Circular structures can support addition and multiplication, as in modular arithmetic (see, e.g., Graham, Knuth, & Patashnik, 1994), and this fact could have legitimized the circular forms as number systems. However, most
participants had probably never encountered modular numbers in their earlier math experience. (In Exp. 3, we asked a similar group of participants at the end of the session whether they had heard of modular numbers. Of the 36 participants, 30 said "no," 3 said "yes," and 3 were unsure or had only minimal experience with these systems.)

Predictors of number-system ratings: Similarity and simplicity

The number forms in Fig. 1 vary in their complexity and in their perceptual similarity to textbook number lines. Could these factors account for the ratings of the number forms as possible number systems? An independent group of participants rated the 25 number forms on these two dimensions, and the means appear in Table 2. For purposes of comparison, the table reverses the complexity scale so that high numbers correspond to the perceptually simplest forms and low numbers to the most complex forms. We will therefore call this factor simplicity rather than complexity in reporting the results. The hypothesis in question is that simpler number forms and forms more similar to a straight-line segment should receive higher ratings as number systems.

Both the mean similarity and simplicity ratings differed significantly across the types of forms [for similarity, $F(4, 124) = 55.83, MSE = 6.72, p < .001$; for simplicity, $F(4, 124) = 31.08, MSE = 4.56, p < .001$]. But the ordering of the means was not the same as in the number-system ratings. Linear number forms, unsurprisingly, received high ratings for both perceptual simplicity and similarity to a line segment, and bilinear forms were also fairly high on these dimensions. However, the step forms obtained the lowest ratings. Multiple-range tests for both perceptual measures showed a significant difference between the linear and bilinear forms, and between the bilinear and the remaining forms, but no further differences. Thus, although participants in the main experiment gave the step forms relatively high ratings as numbers systems, participants in the control group gave them relatively low perceptual ratings.

We might also ask whether the effects of comparability and boundedness persisted in the possible-number-system ratings once similarity and simplicity were controlled. To find out, we regressed the number-system ratings on the similarity and simplicity scores and then reanalyzed the residuals from the regression, using the same type of ANOVA described in the previous section. The results showed that the comparability effect [$F(4, 140) = 5.95, MSE = 9.04, p < .001$], the boundedness effect [$F(4, 140) = 2.88, MSE = 6.84, p = .025$], and their interaction
Our hypothesis was that operation and relational properties will predict number-system ratings better than element properties. Correlations of the weighted property ratings with the possible-number-system ratings appear in Table 1. (To calculate these, we averaged both the weighted property ratings and the possible-number-system ratings over participants and then found Pearson’s $r$ across the 25 number forms.) In line with predictions, the correlations were higher for the relational and operation properties than for the element properties (see the following section for a statistical test). All of the relational properties correlated significantly, but among these items, closure and commutativity produced the weakest correlations. Closure under an arithmetic operation $#$ means that for any two numbers $x$ and $y$, $x \# y$ exists. Participants gave the highest closure ratings to the unbounded forms (those in the first column of Fig. 1) but lower and equivalent ratings to the remaining forms. This pattern of judgments is a reasonable one: For example, addition on a bounded form may mean that certain sums are undefined if they “run off the edge” of the numbers, and the system would therefore fail to be closed under addition. Boundedness, however, did not have a robust effect on the number-system ratings, as we saw earlier, and as a result, closure did not correlate as highly as other relational properties. In the case of commutativity ($x \# y = y \# x$, for an operation $#$), participants gave fairly high ratings to the circular forms, perhaps because the ordering of the numbers is irrelevant for these items. But since the circular forms received very low number-system ratings (see Fig. 2), commutativity was not a strong predictor.

The operation properties in Table 1 all received correlations of .65 or greater, but the correlation for division was somewhat higher than those for the remaining operations. For addition and multiplication, in particular, lower correlations were the result of a pattern similar to the one just discussed in connection with closure: The forms that were unbounded or bounded only below (first and second columns of Fig. 1) received higher weighted property ratings for addition and multiplication than did the remaining forms. This is natural, since participants probably believed that addition and multiplication can continue to produce larger and larger numbers. But this difference failed to capture the number-system ratings, which were generally insensitive to boundedness. The implications of bounding for division may have been less obvious to the participants, and division may have received higher correlations for this reason.

The only element property with a significant correlation to the number-system ratings was density—the existence of an intermediate number between any two numbers. The weighted ratings for density were highest for the (noncircular) single lines, in which intermediate numbers clearly existed, and lowest for the circular forms, in which no ordering was obvious. The remaining number forms received ratings between these. This pattern is roughly consistent with the number-
system ratings, and it accounts for the high correlation. By contrast, most of the remaining element properties—in particular, the presence of a first element, last element, and an infinity of numbers—focused more directly on specific numbers or sets, and were not good predictors of those ratings.

We found earlier that the effects of comparability and boundedness remained significant after controlling for perceptual factors. Did the same effects remain when we control for the mathematical properties? To check this question, we regressed the number-system ratings on composite measures of the operation, relational, and element properties (we describe these measures in the next subsection). The comparability factor was no longer significant in the residuals from this analysis \[ F(4, 140) < 1, MSE = 9.04 \], but the effect of boundedness \[ F(4, 140) = 2.57, MSE = 6.84, p = .040 \] and the interaction of boundedness and comparability \[ F(16, 560) = 2.31, MSE = 1.58, p = .003 \] remained. Thus, the mathematical properties of Table 1 account for essentially all of the effect of comparability. But, as the discussion of the individual properties suggests, these properties predict higher number-system ratings than the data warrant for unbounded forms, and lower ratings than the data warrant for the forms that are bounded above or below.

The same participants in this experiment provided ratings for both the number forms and the number properties. The size of the coefficients could therefore reflect participants’ experience with the task. For this reason, in a second analysis we split the data, using the number-system ratings only from participants who saw the number forms first and the property ratings only from participants who saw the properties first. This analysis was between participants, in that the number system and property data came from separate groups. The resulting correlations, however, showed the same pattern as did the full data set in Table 1. All of the correlations for relational and operation properties continued to be significant at the .01 level. Among the element properties, the first-element, last-element, and successor properties again produced nonsignificant correlations. The only change was that the infinity property was now significant at the .05 level. The stability of these results suggests that the effects are robust over different ways of collecting the data.

Joint predictions. To reduce the number of individual correlations (from the 15 in Table 1), we applied separate principal component analyses to the relational, operation, and element properties. Each analysis took as input the correlations among the weighted property ratings for the properties within a category (e.g., the relational properties). For each category, the analysis returned a first component that represents the best linear composite of the individual properties. For example, one component represented the five properties (associativity, commutativity, etc.) within the relational category. The analysis produced a score for each number form on these relational, operation, and element factors, and we can use these scores to predict the form’s rating as a possible number system. As you might expect from the correlations in Table 1, both the Relational and Operation factors were significant predictors of number forms \( r = .81 \) for the Relational factor and \( r = .83 \) for the Operation factor, \( df = 23, p < .001 \) in both cases). The Element factor, however, failed to correlate significantly with the number-system ratings \( r = .19, df = 23, p = .37 \).

We can get a global measure of how well the math properties predicted the number-system ratings by regressing the ratings against the Relational and Operation factors. For purposes of comparison, the analysis included the similarity-to-a-line measure as an additional predictor. (Similarity and simplicity were themselves highly correlated, \( r = .92, df = 23, p < .001 \). To avoid collinearity, in the analysis we used only the similarity measure.) Thus, the regression contained as the dependent variable the ratings of the 25 number forms as possible number systems, and it contained as independent variables the scores of these forms on the Relational and Operation factors and the mean similarity ratings. The results of this analysis showed that all three variables contributed significantly to the outcome. For the Relational factor, the standardized regression coefficient \( (\beta) \) was .30, \( \beta(21) = 2.13, p = .04 \); for the Operation factor, \( \beta = 0.37, t(21) = 2.59, p = .02 \); and for similarity, \( \beta = 0.38, t(21) = 3.30, p = .003 \). Together, the three variables yielded an \( R^2 \) of .84, \( F(3, 21) = 37.53, MSE = 0.26, p < .001 \). The reliability of the data, computed across participants, was .93, so some systematic variance existed that the model did not account for.

Summary. When participants judged the number forms as possible number systems, their ratings tended to single out possible ambiguities associated with ordering. Circular forms have numbers with indeterminate order, and these forms received uniformly low ratings. Among the noncircular structures, the linear and step forms are completely ordered, as are the two halves of the bilinear forms, and participants gave these items moderate to high ratings. The branching forms, however, are only partially ordered, and they received lower ratings. Participants paid less attention to whether the forms were bounded or unbounded: Ratings were not significantly higher for the unbounded items than for those bounded above or below. These priorities also showed up in the correlations with the mathematical properties of Table 1: Properties that tapped ordering and arithmetic tended to be good predictors of...
the number-system ratings, and the effects of these math properties were independent of purely perceptual factors, such as similarity to a straight line.

Although some of the mathematical properties in Table 1 successfully predicted which forms participants considered to be possible number systems, they were not necessarily the properties that people themselves consider most important. Perhaps people know of an element property that they believe is crucial to number systems but that was not among those in Experiment 1. If so, then the results might have understated the importance of the element properties. Experiment 2 investigated this possibility by asking participants to generate properties that they believed would discriminate number systems from more arbitrary sets of numbers. These features may provide further insights into why nonmathematicians consider some number forms more likely than others to be number systems.

Experiment 2: Generated properties as predictors of possible number systems

The main part of this study followed the procedure of Experiment 1: Participants rated the number forms in Fig. 1 according to whether they could serve as possible number systems. They also rated a set of properties for their importance in number systems. However, the new experiment included two main changes. First, the participants in a preliminary study wrote down properties that they believed number systems incorporate. We asked them to imagine that they were anthropologists studying an isolated culture that had developed ways of talking about numbers. Their goal was to determine whether the natives' numbers formed a number system by asking native mathematicians about their numbers. Participants wrote down ten properties that they would investigate to determine whether the numbers formed a number system. These properties served as the basis for those that appeared in the main part of the experiment, replacing the properties in Table 1.

A second change in this experiment was that the number forms showed the position of zero. The forms in Experiment 1 had contained no labeled numbers, and so the positions of the positive and negative numbers were uncertain. Participants in the preliminary experiment produced some properties that mentioned the positive and negative numbers, and we therefore included a zero point to mark this distinction. An additional advantage of displaying the zero points was that they highlighted one of the form's elements. Perhaps the advantage of operation and relational properties over element properties in Experiment 1 was due to the absence of any explicit elements on the number forms. The presence of zero in this experiment therefore provided a stronger test of our hypothesis. The resulting forms were similar to those of Fig. 1. As is shown in the figure, each bilinear form contained two zeros (represented by 0 and $0'$), to go along with the idea that the form included two sets of incommensurable numbers. Branching forms contained just a single zero, on the forms' main "trunk."

Method

Two groups of participants took part in the study. One group produced properties that would help identify a number system. A second group evaluated a subset of these properties for their importance in number systems, indicated which of the number forms in Fig. 1 these properties applied to, and rated the number forms as possible number systems.

Property generation Participants received a two-page booklet containing instructions on the first page and space for writing on the second. The instructions described a scenario in which an isolated group of native people had developed numbers, "though not necessarily numbers with which you are familiar." The instructions continued:

imagine that your goal is to determine whether the native's numbers form some type of number system. You've learned about several different number systems in school, such as the natural numbers, the integers, the rational numbers, the real numbers, the complex numbers, and so on. What you would like to find out is whether these people also have a number system. But you don't want to confine yourself to familiar number systems, such as those just mentioned. You want to hold open the possibility that the natives may have discovered a new number system.

Participants were to choose properties about which they would ask native mathematicians in order to determine whether their numbers formed a possible number system. The following page contained spaces for ten such properties. The participants took 10–15 min to complete the task, and it appeared in the same session as activities for two unrelated studies.

To select the properties for the rating task from those the participants had produced, we classified the properties according to the categories of Experiment 1 (relations, elements, and operation), but added two further categories to accommodate some remaining properties. Relations accounted for 24 % of the tokens, elements for 24 %, and operations for 13 %. Table 3 contains examples of these items. An additional 20 % of the properties concerned possible uses of a number system—for example, "Can the numbers be used to measure things?" and "Is it possible to plot the numbers?" Another small set of property tokens (4 %) mentioned notation (e.g., "Do the numbers have a symbolic representation?"). The remaining 15 % of tokens were classified as "other" (e.g., "Integrated with logic," "Is it different from language?"). We
Table 3 Stimulus properties in Experiment 2, with mean ratings of importance to number systems and correlations of the weighted property ratings to the ratings of number forms as possible number systems

<table>
<thead>
<tr>
<th>Relations</th>
<th>Mean Rated Importance</th>
<th>Correlation With Number-System Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ordering: The numbers in S have a constant order.</td>
<td>5.80</td>
<td>.91**</td>
</tr>
<tr>
<td>Constant difference: Any two consecutive numbers in S have a constant difference.</td>
<td>5.06</td>
<td>.90**</td>
</tr>
<tr>
<td>Closure: Any two numbers in S can be combined to yield a number.</td>
<td>5.74</td>
<td>.59**</td>
</tr>
<tr>
<td>Constant result: A given operation on any two numbers in S always produces the same result.</td>
<td>5.40</td>
<td>.72**</td>
</tr>
<tr>
<td>Constant value: Any number in S has a constant value.</td>
<td>6.17</td>
<td>.86**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Elements</th>
<th>Mean Rated Importance</th>
<th>Correlation With Number-System Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero: S contains zero.</td>
<td>4.34</td>
<td>.13</td>
</tr>
<tr>
<td>Negatives: S contains the negative numbers.</td>
<td>3.77</td>
<td>.46*</td>
</tr>
<tr>
<td>Positives: S contains the positive numbers.</td>
<td>3.68</td>
<td>.56**</td>
</tr>
<tr>
<td>Infinity: S contains an infinity of numbers.</td>
<td>5.03</td>
<td>.45*</td>
</tr>
<tr>
<td>Integers: S contains the whole numbers.</td>
<td>3.48</td>
<td>.84**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Operations</th>
<th>Mean Rated Importance</th>
<th>Correlation With Number-System Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addition: Addition can be defined on S.</td>
<td>5.91</td>
<td>.67**</td>
</tr>
<tr>
<td>Multiplication: Multiplication can be defined on S.</td>
<td>5.97</td>
<td>.86**</td>
</tr>
<tr>
<td>Subtraction: Subtraction can be defined on S.</td>
<td>5.94</td>
<td>.85**</td>
</tr>
<tr>
<td>Division: Division can be defined on S.</td>
<td>5.14</td>
<td>.91**</td>
</tr>
<tr>
<td>Operation: Some mathematical operation can be defined on S.</td>
<td>6.37</td>
<td>.96**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Uses</th>
<th>Mean Rated Importance</th>
<th>Correlation With Number-System Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counting: The numbers in S can be used for counting.</td>
<td>4.11</td>
<td>.97**</td>
</tr>
<tr>
<td>Measuring: The numbers in S can be used for measuring.</td>
<td>4.11</td>
<td>.90**</td>
</tr>
<tr>
<td>Graphing: The numbers in S can be graphed.</td>
<td>5.06</td>
<td>.80**</td>
</tr>
</tbody>
</table>

*p < .05, **p < .01, df = 23*

then grouped individual properties if they shared the same core meaning (e.g., “Can you add the numbers together to get another one?” and “Can the numbers be added?”). For the main experiment, we selected the five most frequent properties from each of the “relations,” “elements,” and “operations” categories. At least two participants had listed each of these properties. We also included the three most frequent use-based properties—the only such properties to receive more than one mention. None of the properties in the “notation” or “other” categories had been listed by more than one participant, and these properties were dropped from the final set. The resulting properties, grouped by categories, appear in Table 3.

Number form and property ratings: The procedure in the main experiment followed that of Experiment 1. The participants rated the number forms in Fig. 1 according to whether they were possible number systems. The instructions were the same as in the earlier experiment, except that they mentioned, “a zero point will appear on each line with a tick mark to indicate its position. If there are two lines, there may be two zero points (one marked 0 and the other 0).” In the property-rating part of the experiment, participants saw each of the 18 properties from Table 3, and they indicated on a 1–8 scale how important the properties were to possible number systems. After completing these property ratings, the participants received a booklet, each page of which contained one of the properties at the top and an array of the 25 number forms beneath. Participants circled “yes” or “no” beneath each form to indicate whether the property was true or false of the form. Sixteen participants rated the number forms, rated the properties, and then decided which properties applied to which forms. Nineteen participants rated the properties, decided on the property–form pairings, and then rated the forms.

Participants: A group of 14 participants generated properties in this study, and 35 performed the number form and property ratings. These participants were from the same pool as those in Experiment 1, but none had taken part in the earlier experiment.

Results and discussion

Ratings of number forms as possible number systems: The ratings of the number forms appear in Fig. 3, and once again are highest for the linear forms (M = 6.24) and next highest for the bilinear ones (M = 5.50). The step forms (M = 3.74), forward-branching forms (M = 4.19), and backward-branching forms (M = 4.15) received lower ratings. The difference among these forms was reliable [F(4, 132) = 35.98, MSE = 5.67, p < .001]. A multiple-range test showed differences between the linear and bilinear forms, and between the bilinear forms and the remaining ones, but no further significant differences. Participants who rated the forms before rating the properties showed a wider range in their possible-number-system ratings than did those who did these tasks in the opposite order [F(4, 132) = 4.37, MSE = 5.67, p = .002], but the relative orderings of the means were the same for both groups.

These number-system ratings were similar to those from Experiment 1 (see Table 2); however, the step forms received lower ratings in this study (M = 5.24 in Exp. 1 vs. M = 3.74 here). The forms in the two experiments differed just in the
appearance of the zero point. The presence of a zero on the upper (but not the lower) part of the step forms (see Fig. 1) may have conveyed to participants a possible difference in meaning between the upper and lower segments, and this implication may have lowered the ratings for these items, relative to those in Experiment 1. Written comments from the participants at the end of the session suggested that some of them interpreted the forms as functions. According to this interpretation, vertical lines represent multiple values (as they would on the $y$-axis of a Cartesian plane) for a given $x$-axis number. These participants gave lower number-system ratings to all forms that contained vertical lines, resulting in lower ratings for the step and branching items. For example, one participant wrote, “I rated the diagrams by considering whether or not there were any lines occupying the same vertical space (like a function in mathematics). If there was any vertical overlap, I counted it as not a number system because then there could be one number with two values.” We addressed this issue in Experiment 3 by indicating explicitly the positions of the numbers on the number forms.

As in Experiment 1, the effect of boundedness—the column factor in Fig. 1—was confined to the circular forms. Although this factor produced a significant effect [$F(4, 132) = 32.10$, $MSE = 5.20$, $p < .001$], a multiple-range test identified as reliable only the difference between the circular forms and the remaining items. Figure 3 shows that the circular forms all received low ratings, and the difference between circular forms and the other boundedness conditions produced an interaction between the Boundedness and Comparability factors [$F(16, 528) = 5.51$, $MSE = 1.33$, $p < .001$]—a result that echoes that of Experiment 1.

Math properties The properties that participants rated in this experiment were ones that the preliminary group had generated, and for this reason, they may have seemed more natural than the textbook properties of Experiment 1. However, the overall mean rating of importance ($M = 5.06$) was no higher than that of the earlier experiment ($M = 5.46$), and the ordering for the categories of operation, relational, and element properties was similar. The importance rating for each property appears in Table 3. Operation properties, such as the ability to support addition and multiplication, produced high ratings ($M = 5.87$), as had the comparable properties in Experiment 1. Participants also gave high ratings to relational properties ($M = 5.63$), most of which stressed the stability of the number system—for example, having a constant order of numbers and having a constant difference between adjacent numbers. Participants gave lower ratings to the existence of specific elements ($M = 4.06$) and to possible uses of the numbers, such as in counting and measuring ($M = 4.43$). The main effect of property type was significant in this study [$F(3, 99) = 21.63$, $MSE = 6.15$, $p < .001$], and a multiple-range test confirmed the difference between the relational and operation properties, on the one hand, and the element and use-based properties on the other. No further differences were significant among these property types, and task order (property-rating first vs. number-form rating first) did not interact with the property-type effect, $F(3, 99) = 1.84$, $MSE = 6.15$, $p = .14$. 

![Fig. 3 Mean ratings of number forms as possible number systems in Experiment 2. Circles represent linear forms; squares, bilinear forms; diamonds, step forms; upward triangles, forward-branching forms; and downward triangles, backward-branching forms. Error bars represent ±1 standard error of the mean, as estimated from the mean squared error in the reported analysis of variance](image)
We could also check whether these new properties predicted the forms’ number-system ratings. A participant’s rating of a property’s importance was multiplied by +1 if the participant said that it applied to a number form, and by −1 if the participant said that it did not, following the procedure of Experiment 1. Correlations between the means of these weighted properties and the means of the possible-number-system ratings appear in the last column of Table 3. These correlations are somewhat higher than those in Experiment 1, perhaps because of the greater familiarity of the properties. The only nonsignificant correlation was whether the number system contained a zero. Because all number forms included a zero, this property did not discriminate among them.

In general, higher correlations appeared for the relational and operation properties than for the element properties—the same pattern that had occurred in the first experiment. Use-based properties also received high correlations. As we mentioned, most of the relational properties focused on the stability of the number system—whether the forms guaranteed a constant ordering of the numbers, a constant difference between them, or constant results for arithmetic. Weighted property ratings for these characteristics were highest for the (noncircular) linear and bilinear forms, and lowest for the circular forms. This pattern produced high correlations with the possible-number-system ratings (see Table 3). By contrast, the closure property (here stated as “Any two numbers in S can be combined to yield a number”) produced lower correlations, as it had in Experiment 1. The operation properties—for example, the property that division can be defined on the numbers—all correlated fairly highly with the number-system ratings.

Although the element properties were significant predictors in this experiment, their correlations tended to be smaller than those from the other categories. Within this group, the integer property (“S contains the whole numbers”) was the only one whose correlation could compete with those of the other property categories, and the reason for its success is unclear.

Despite the relatively low rated importance of the use-based properties, they correlated quite highly with the number-system ratings. Participants’ comments suggested that they considered the ability to use a set of numbers for counting and other purposes to be one test of whether a set of numbers qualified as a number system, and they were puzzled by how a circular system or a branching system could achieve these goals.

Joint predictions To consolidate the properties, we extracted the first principal component for each of the four property categories (see the Results and Discussion section of Exp. 1). We then correlated the scores on these components with the possible-number-system ratings. In accord with the impression from Table 3, correlations were high for the relational $[r(23) = .87, p < .001]$, operation $[r(23) = .90, p < .001]$, and use-based $[r(23) = .93, p < .001]$ properties, but somewhat lower for the element properties $[r(23) = .68, p < .001]$. The relational, operation, and use-based scores were themselves highly correlated; the smallest of these intercorrelations was .92.

In Experiment 1, we found that the mathematical properties predicted the number-system ratings in a way that was independent of the perceptual similarity of the number forms to a straight line. We did not collect similarity ratings in the present experiment: The zero-points on the number forms would have revealed to participants the mathematical nature of the forms, and we wanted the similarity ratings to measure purely perceptual features. Apart from the zeros, though, the number forms were the same in the two studies, and we could compare the earlier similarity ratings to the new set of mathematical properties as predictors in the present experiment. Because of the intercorrelations among the properties, just mentioned, we used only the scores for the operation properties and the similarity ratings as independent variables in a regression.

The results produced significant coefficients for both variables $[\beta = 0.63, t(22) = 4.29, p < .001]$, for similarity, and $[\beta = 0.34, t(22) = 2.32, p = .03]$, for the operation factor. The model produced an $R^2$ of .69, $F(2, 22) = 90.64$, $MSE = 0.12$, $p < .001$. Substituting either the use-based scores or the relational scores in place of the operation scores produced models that did approximately as well as the one just reported. As in Experiment 1, these models left some systematic variance unaccounted for, since the reliability of the data was .96.

**Experiment 3: Discrete number systems**

The number forms of Experiment 1 did not indicate the positions of individual numbers, and those of Experiment 2 indicated only the position of zero. These forms encouraged participants to regard the represented set of numbers as continuous. Continuity, however, may have affected how the participants evaluated the forms as possible number systems (see Alibali, Bassok, Solomon, Syc, & Goldin-Meadow, 1999, for effects of discrete vs. continuous variation on problem solving). For example, continuity ensured that all forms had an infinite number of elements, even if the forms were bounded both above and below. This fact may explain why the boundedness factor had only a minimal impact on the possible-number-system ratings. Similarly, the absence of symbols for individual elements may help explain why the element properties proved weaker predictors of the number-system ratings than the relational or operation properties. In the present experiment, we examined the effect of using discretely many rather than continuously many numbers.

The change from continuous to discrete numbers also allowed us to avoid a potential misinterpretation of the number
forms in Experiment 2. As we mentioned, some participants regarded the vertical line segments in the step and branching forms as representing elements along an implicit $y$-axis. To help prevent this interpretation in the present experiment, we used tick marks to indicate individual numbers, and we informed participants that numbers occurred only at the positions of the ticks. Figure 4 shows the resulting number forms. The ticks on the parallel lines for the bilinear and branching forms were offset horizontally to discourage the impression that some positions on the forms had multiple “values,” one above the other.

Method

Apart from changing the forms (and accompanying instructions) to include tick marks, this experiment duplicated the method of Experiment 1. The instructions for the possible-number-system ratings told participants that they would be seeing diagrams containing “a series of numbers (one for each tick mark on the line), ordered in the direction of the arrows from smallest to largest. The only numbers are at the positions marked by the ticks.” To help prevent participants from regarding the lines as graphs on a co-ordinate plane, the instructions also stated, “These are number lines, NOT functions. Please do NOT think of the lines [as] part of a co-ordinate system with an $x$- and $y$-axis. The only form of ordering for the numbers is provided by the arrows.”

The property-rating task employed the properties and instructions of Experiment 1. As in the earlier experiments, participants also indicated for each property whether it was true or false of the individual diagrams. The procedure was the same as in Experiment 1, but the instructions included the information about tick marks and the warning not to interpret the lines as graphs.

The participants were 36 students from the same pool as those in Experiments 1 and 2 (though none had taken part in those experiments). Half of the participants rated the number forms, then rated the properties, and finally indicated which properties were true of which forms. The remaining participants rated the properties, indicated which properties were true of which forms, and then rated the forms.

Results and discussion

The results in this study were generally similar to those of Experiment 1, but with a few differences due to the new discrete number lines. A preliminary analysis showed no main effect of the order in which participants carried out the tasks and no interaction of task order with the remaining variables. We will therefore collapse over order in presenting the results.
Ratings of the number forms as possible number systems. Linear forms again received the highest ratings \((M = 6.56\) on the 8-point scale), followed by the bilinear \((M = 5.72\) and step forms \((M = 5.17\). The branching forms drew lower ratings \((M = 4.56\) for the forward-branching items and \(M = 4.38\) for the backward-branching items). The difference among these means was reliable, \(F(4, 140) = 18.05, MSE = 7.87, p < .001\). Figure 5 displays the data for all 25 forms.

One reason for the use of discrete forms in this experiment was to see whether exhibiting the positions of individual numbers would clarify the meaning of the diagrams. We suggested that participants in Experiment 2 might have misunderstood the vertical segments as if the forms were plots of functions. Such an interpretation would tend to lower the ratings for the step forms, which have several of these vertical segments. By including tick marks for the numbers (and stressing that the forms were not functions), we hoped to make it clear that no numbers occurred at the positions of the vertical lines. The results showed that mean ratings for the step forms were, in fact, noticeably higher here than in Experiment 2 \((M = 5.17\) vs. 3.74). Of course, each of the branching forms also had a vertical segment, so the new format could potentially benefit these items too. In fact, a small change in this direction appeared in the data, with mean ratings for the branching items increasing from 4.17 in the second experiment to 4.47 in this one. The branching forms, however, also created ambiguities about arithmetic, as we noted earlier, and this property could have kept participants from giving the branching items high marks as number systems.

A second motive for the discrete forms was to find out whether discreteness would enhance the influence of boundedness, and the results suggested that this, too, was the case. Figure 5 shows that the circular forms again received lower ratings than did the other types of bounding. As in Experiments 1 and 2, the boundedness main effect and the interaction with the comparability variable were both significant \([\text{for the main effect, } F(4, 140) = 23.44, MSE = 7.62, p < .001; \text{ for the interaction, } F(16, 560) = 2.38, MSE = 1.40, p = .002}\]. A multiple-range test (REGW with \(\alpha = .05\)) showed the usual difference between the circular forms and the noncircular ones. In this experiment, however, the difference among the noncircular forms was somewhat larger than in the previous experiments, with the unbounded forms (first column in Fig. 4) scoring higher than the bounded ones (fourth column). (No further difference in bounding among the noncircular forms was significant.) Our instructions about the tick marks implied that the bounded forms and the circular forms contained only finitely many numbers, and participants may have wondered whether a finite set could serve as a legitimate number system. Their written comments after the experiment sometimes mentioned infinity as a criterion. According to one participant, for example, “The number systems that were infinite seemed more likely to me, because as a system, it seems that there should be an infinite amount of
numbers.” Another participant asserted, “Systems that were not infinite could not be number systems.”

**Math properties** Participants in this experiment rated the mathematical properties from Experiment 1 according to their importance for number systems. These ratings appear in Table 4, and they agree quite well with those of Experiment 1 (see Table 1). The highest ratings again went to the operation properties ($M = 6.36$), followed by the relational properties ($M = 5.38$) and the element properties ($M = 5.10$), $F(2, 68) = 9.60$, $MSE = 8.10$, $p = .0002$. The correlation between the ratings of the individual properties from Experiments 1 and 3 was $r(13) = .95$, $p < .001$.

To find out whether these mathematical properties could predict the ratings of the number forms, we weighted the properties, following the procedure that we had used in the earlier experiments. For each property, we then correlated the weighted ratings with the ratings of the number forms. The correlations also appear in Table 4, and they are consistent with the corresponding figures from Experiment 1 that appear in Table 1. We once again found that the relational and operation properties produced noticeably higher correlations than did the element properties. The median correlations for the relational and operation properties were .81 and .87, respectively, but that of the element properties was only .25.

The absence of tick marks (symbols for individual numbers) in the forms of Experiment 1, then, was probably not the reason why participants gave relatively low ratings to the element properties, nor why the element properties failed to predict the number-system ratings in that experiment.

**Joint predictions** We extracted single factors to represent the operation, relational, and element properties using principal component analysis, as in the previous experiments. Correlations between these factors and the number-system ratings were high for operation properties [$r(23) = .95$, $p < .001$] and relational properties [$r(23) = .93$, $p < .001$], but near zero for the element properties [$r(23) = -.03$, $p > .10$]. This pattern is a more extreme version of the one that we had come to expect from Experiments 1 and 2.

To compare the effects of these mathematical properties to that of the similarity to a straight line, we performed regressions with these factors as independent variables and the number-system ratings as the dependent variable. (For these purposes, we used the similarity scores from Exp. 1, for the reasons discussed earlier.) Because the Operation and Relational factors were highly correlated [$r(23) = .96$], we included the operation and similarity scores in one model and the relational and similarity scores in another. For the first of these regressions, both coefficients were significant [for the Operation factor, $\beta = 0.74$, $t(22) = 11.85$, $p < .001$, and for similarity, $\beta = 0.30$, $t(22) = 4.85$, $p < .001$], with $R^2 = .95$, $F(2, 22) = 230.76$, $MSE = 0.05$, $p < .001$. The $R^2$ value in this case was nearly equal to the reliability of the data. The second regression produced similar results [for the Relational factor, $\beta = 0.73$, $t(22) = 6.93$, $p < .001$, and for similarity, $\beta = 0.26$, $t(22) = 2.44$, $p = .02$], with $R^2 = .89$, $F(2, 22) = 93.20$, $MSE = 0.12$, $p < .001$. Although simple perceptual similarity may be responsible for some aspects of

| Table 4 Stimulus properties in Experiment 3, with mean ratings of importance to number systems and correlations of the weighted property ratings to the ratings of number forms as possible number systems |
|---------------------------------|----------------|----------------|
| Relations                      | Mean Rated Importance | Correlation With Number-System Ratings |
| Associativity: If $\#$ is an operation (for example, addition or subtraction) defined on $S$, $x \# (y \# z) = (x \# y) \# z$, for any $x, y, z \in S$. | 5.03 | .89** |
| Commutativity: If $\#$ is an operation (for example, addition or subtraction) defined on $S$, $x \# y = y \# x$, for any $x$ and $y \in S$. | 4.83 | .81** |
| Closure: If $\#$ is an operation (e.g., addition or subtraction) defined on $S$, $x \# y$ is also in $S$, for any $x$ and $y \in S$. | 5.86 | .44* |
| Trichotomy: For any $x$ and $y$ in $S$, either $x < y$ or $y < x$ or $x = y$. | 6.97 | .77** |
| Inequality preserved: If $\#$ is an operation (e.g., addition or subtraction) defined on $S$ and $x > y$, then $z \# x > z \# y$, for any $x, y, z \in S$. | 4.19 | .95** |
| Elements                       |                |                |
| First element: $S$ has a first element. | 4.94 | .16 |
| Last element: $S$ has a last element. | 3.53 | .25 |
| Infinity: There are an infinite number of elements in $S$. | 5.58 | .43* |
| Successor: For any $x \in S$, there is a number immediately following $x$. | 6.06 | .01 |
| Density: For any $x$ and $y \in S$, there is a number $w$ between them: $x < w < y$ or $y < w < x$. | 5.42 | .86** |
| Operations                     |                |                |
| Addition: Addition can be defined on $S$. | 6.64 | .87** |
| Multiplication: Multiplication can be defined on $S$. | 6.39 | .84** |
| Subtraction: Subtraction can be defined on $S$. | 6.53 | .87** |
| Division: Division can be defined on $S$. | 5.75 | .94** |
| Operation: Some mathematical operation can be defined on $S$. | 6.47 | .95** |

*p < .05, ** p < .01, df = 23
the findings, the participants also took into account more formal properties in evaluating the number forms.

**General discussion**

Participants in these experiments decided whether each of a set of “number lines” could represent a number system. Some of these number forms, however, were unusual. Besides simple line segments, we included forms with two independent sets of numbers (bilinear items), forms that partially ordered the numbers (branching items), and single lines with a zigzag shape (step forms). In Experiment 1, the forms contained no tick marks for individual numbers, but in Experiment 2, they showed the position of zero (as in Fig. 1), allowing for a distinction between positive and negative numbers. The forms in Experiment 3 had additional ticks for other numbers. The forms were continuous in Experiments 1 and 2, but discrete in Experiment 3. In all studies, participants considered single lines to be clear examples of number systems, the bilinear forms to be of intermediate status, and the branching forms to be of lower status. Results for the step forms differed among the experiments: They were equivalent to the bilinear forms in Experiments 1 and 3, but equivalent to the branching forms in Experiment 2. The number forms also differed in their boundaries: They were either unbounded (extending infinitely in both positive and negative directions), bounded only above, bounded only below, bounded both above and below, or circular. Circular forms received low ratings as number systems in all experiments. Other forms of bounding had generally higher but similar ratings, except in Experiment 3, where participants gave lower ratings to forms with a finite number of elements.

These results suggest that participants evaluated the number forms in terms of the internal ordering of the numbers rather than of the presence of particular numbers or boundaries. Having all the positive or all the negative numbers, for example, mattered less than whether the numbers were completely ordered. Although we considered circular arrangements as a type of bounding for purposes of the analysis, these forms make the ordering of numbers arbitrary, and this effect on ordering may explain why they consistently received low ratings.

These conclusions go along with the results on mathematical properties. In Experiments 1 and 3, participants rated a set of textbook properties (e.g., associativity) for their importance to number systems. In Experiment 2, they did the same for a set of properties that an independent group had produced as being relevant to number systems. Most properties that specified particular numbers (e.g., a first number, positive numbers, or negative numbers) did less well as predictors of the possible-number-system ratings than did properties that tapped relations among the numbers, their ability to support arithmetic operations, or their usefulness in counting or measuring. These results suggest some degree of abstractness in people’s concept of a number system. The college students who took part in these studies had experience with number systems of different types, and this experience may have enabled them to generalize over features that were local to particular systems, such as the natural numbers. However, other aspects of the data hinted at ways in which the students’ concepts departed from those of mathematicians. The rest of this discussion tries to clarify this difference by relating it to experts’ theories of number systems.

**What are number systems?**

Mathematicians universally recognize the natural numbers, integers, rationals, reals, and complex numbers as being number systems (Feferman, 1964). But, perhaps surprisingly, no formal definition of number system seems to exist. To understand this concept, we need to adopt an indirect approach—similar to the approach that cognitive psychologists adopt for studying concepts of most natural kinds and artifacts—looking at the role that number systems play in mathematics.

One place to start is set theory, since a central task in this field is to formalize the properties of basic number systems and to construct further systems from simpler ones. For example, set theory (e.g., Enderton, 1977; Hamilton, 1982) has shown how to define the integers as sets of pairs of natural numbers: For example, –3 is defined as the set of pairs of natural numbers \(<0, 3>, <1, 4>, <2, 5>, \ldots\), whose difference is –3. Similarly, these texts define the rational numbers as sets of pairs of integers: For example, \(-1/2\) is \(<-1, 2>, <1, -2>, <2, 4>, <2, -4>, \ldots\), which all yield \(-1/2\) when the first element is divided by the second. Given these definitions, theorems establish that arithmetic operations (e.g., + and *) are well defined over the sets representing the new numbers (where well defined means that arithmetic comes out the same, no matter which elements we choose from the representing sets). The texts then prove additional theorems governing other properties of the new system: For example, associativity holds under addition and multiplication; identity elements and inverses exist; and so on.

These constructions use previously defined number systems (e.g., the natural numbers) to produce new ones (e.g., the integers). But something like the same pattern appears when formalizing the initial systems. In the case of the natural numbers, the Dedekind–Peano axioms begin with a single starting number 0, a successor relation, and a closure (or induction) principle. Subsequent proofs show how to define addition and multiplication and to verify their properties. Of course, the natural numbers (and systems like the rationals that can be constructed from them) were familiar to mathematicians prior to their formal definitions, and the definitions aimed to capture their known properties. However, the
development of new number systems also proceeds in the same way. In Knuth’s (1974) mathematical novella, *Surreal Numbers*, the young protagonists explore a new number system (from Conway, 2001) beginning with 0 and a definition of ≤. They then prove properties of the ordering relation, define addition, and confirm properties of addition such as associativity.

Students’ conceptions of number systems

In these terms, the present results suggest that college students are sensitive to many of the right features, but may have an overly narrow idea of the key concept. On the one hand, participants found arithmetic operations to be central to number systems. The operation properties, as a group, received the highest ratings in all experiments. The specific property “Some mathematical operation can be defined on $S$” received the highest rating of all of the properties in Experiment 2, and relatively high ratings in Experiments 1 and 3. It was also a significant predictor of the number-system ratings in each study.

But, on the other hand, participants resisted the idea of a number system that violates complete ordering. In some cases, this is because the violations suggested that the arithmetic operations would produce ambiguous results. Poor marks for the branching forms may reflect this kind of thinking. One participant in Experiment 2, for example, explained his or her ratings in the following way in written comments following the session: “For the diagrams that split I did not consider them number systems, because that would require some type of function to yield 2 responses whereas a number system should be consecutive.” Another participant wrote, “Systems that diverged into two or more lines indicate that there is a possibility that mathematical functions would not have a standard answer (i.e., one operation on two numbers could result in either of two solutions).” These are perfectly understandable evaluations. However, participants also gave low ratings to circular systems, despite the fact that such systems can support addition and multiplication, associativity, and other properties. Similarly, participants assessed density and trichotomy quite positively, in both their rated importance and their correlation with the number-system ratings, even though some standard number systems violate such properties. For example, the natural numbers and the integers are not dense (e.g., no integer is between 0 and 1), and the complex numbers do not obey trichotomy, as mentioned earlier.

A structuralist account of number systems

We can begin to make sense of these findings from the standpoint of a structuralist theory of mathematics (Resnik, 1997; Shapiro, 1997). According to this view, a mathematical object, such as a particular number or set, depends for its existence on its place within a system. The natural number 0, for example, is whatever plays the role of the initial element in the successor sequence, as defined by the axioms for this system. The natural number 2 is whatever plays the role of the successor of the successor of that initial element. From this standpoint, elements derive from relations and functions, and thus the latter take precedence over the former in specifying the underlying nature of mathematics.

We might view the results of these experiments as supporting a structuralist view by showing that people’s concept of number system also privileges relations and operations over elements. As a psychological proposal, however, giving priority to math relations and operations runs counter to standard assumptions. The usual notion, at least with respect to the natural numbers, is that math concepts get their meaning from their connection to physical objects (see Rips, 2011, for a review of these theories). The concept of 3, for example, is the concept of all three-element sets of objects, such as three bears and three stooges. According to this view, the concept of 3 has a meaning that is independent of that of 2 (the concept of all two-element sets) or that of 4 (the concept of all four-element sets). Much the same assumption appears in educational contexts, where learning the meaning of 3 is learning how to apply a counting technique to determine that a set contains three objects. According to such theories, children who have not yet learned to count (and natives of cultures who do not have a technique for counting) do not have the concept of individual natural numbers—at least not beyond the range of objects that they can track in immediate memory. By contrast, the relations among the natural numbers are sometimes regarded as mere placeholders rather than as being constitutive of the numerals’ meaning (but see Leslie, Gallistel, & Gelman, 2008, for an opposing view).

An objection based on sparseness of the representations

Although the structuralist interpretation of these experiments is an inviting one, several aspects of the present data suggest qualifications. Perhaps the most obvious of these is that the number forms in these experiments contained little more than bare lines, minimizing the representation of particular elements. The forms in Experiment 1 showed no specific numbers, and those in Experiment 2 showed only the position of zero (see Fig. 1). Thus, the fact that relational and operation properties were better predictors than elements could simply have been due to the nature of the displays. Textbook number lines are normally less stingy than ours in showing individual numbers or number locations, so perhaps the superiority of the relational and operation properties over the elements was due to the lack of information about the latter.

The number forms in these experiments place some limits on what we can learn from them about people’s concepts. The objection to bare lines, however, has to cope with several findings. First, the addition of ticks for numbers in Experiment 3 did not produce any further support for the
element properties. Second, the element properties themselves received lower ratings for their importance to number systems than did the operation and relational properties. People apparently believe that specific elements are less important than operations or relations, even when they evaluate these properties directly. Of course, the lower ratings for elements could be due to a biased sampling of these properties. Perhaps the elements that we chose as stimulus properties in Experiments 1 and 3 were not the most important ones for evaluating number systems, at least in the participants’ estimation. In Experiment 2, however, the properties came from the preliminary study. Here, participants produced the properties they thought would determine whether a set of numbers was a number system, and we selected the most frequent of these to include in the main experiment.

Finally, a look at the weighted property ratings shows that participants were able to discriminate the number forms that embodied the element properties from those that did not, despite the forms’ sparseness. For example, the weighted ratings for the property “S contains the positive numbers” were higher for the number forms that were unbounded and bounded below than for the remaining forms in Experiment 2. Similarly, weighted ratings for the element property “S contains the negative numbers” were uniformly higher for the forms that were unbounded and bounded above. These differences suggest that participants were able to make sensible decisions about the presence of specific elements, even with stripped-down number forms.

An objection based on usefulness The structuralist view of these experiments is open to criticism on a second front. Experiment 2 showed that the properties “can be used for counting” and “can be used for measuring” correlated very highly with the possible-number-system ratings (see Table 3). If people think of number systems as being defined by their internal organization, why are these use-based properties such good predictors of whether a number form represents a possible number system? Branching or circularity in the number forms makes it difficult to see how people could use these forms to enumerate objects in the normal way. Counting with these forms would presumably violate standard counting principles—for example, the principle that numerals should be paired with the objects to be counted on a one-to-one basis (Gelman & Gallistel, 1978). Perhaps the use-based properties are what drive participants’ decisions about number systems, with the relational and operation properties entering these decisions mainly in supporting these uses.

But although the use-based properties correlated well with the number-system ratings, they did not get high ratings on their own for their importance to number systems. Their mean rating was about the same as that for element properties, and significantly lower than those for relational or operation properties, as we noted earlier (see Table 3). The weighted use-based properties may predict the number-system ratings not because of their own importance, but because they presuppose other aspects of these systems that are important. Both counting and measuring require that numbers be ordered in specific ways and that they obey basic arithmetic operations (e.g., Krantz, Luce, Suppes, & Tversky, 1971). This is not to deny that people may wonder about a potential number system that cannot be used for external purposes. But no evidence from the present experiments exists to show that these considerations take precedence over a system’s internal arrangements. Finally, the usefulness of a number system does not imply that the numbers get their meaning from their use. The structuralist account is perfectly compatible with applications of number systems to practical or scientific tasks.

We cannot hope to settle on a structuralist theory of math concepts only on the basis of the experiments reported here. Still, the results suggest that people emphasized facts about the numbers’ internal organization in the tasks that they performed. The possible-number-system ratings were more sensitive to the comparability of the numbers under less than (<) and greater than (>) than to the existence of a particular set of numbers. Participants also rated the presence of relations and operations as being higher in importance for number systems than the presence of specific elements, and the former properties had higher correlations with the ratings of the number forms than did the latter. Internal structure takes center stage when people decide what a possible number system might be.

Conclusions

College students’ beliefs about number systems are correctly attuned to the way that such systems accommodate arithmetic and to the relations that arithmetic imposes. They know, for example, that these operations are functions that produce a unique result (see Table 3 and the comments from participants quoted earlier). They are properly less concerned about the existence of particular numbers or number ranges.

However, the results also point to ways in which even college students’ concepts of number systems may be too narrow. The importance that they attach to trichotomy (for any $x$ and $y$, $x < y$, $x > y$, or $x = y$) and similar properties suggests that they are still bound to linearity. This reliance on linear ordering might make it difficult for them to understand systems like the complex numbers that violate such constraints. This emphasis recalls the ease of linear ordering relative to partial ordering in other tasks (see, e.g., De Soto, 1961; Hayes-Roth & Hayes-Roth, 1975; Moeser, 1979) and may reflect similar cognitive constraints. We have suggested that the results of these experiments are consistent with a structuralist approach in which the meaning of a particular number is determined, not by its denoting an individual object or set, but instead by its role in an organized framework. If this
is correct, then attempts to teach students new numbers systems may be more successful if they focus on properties of the system as a whole. Instruction should provide help when these properties become difficult to envision, and it should explain why old properties no longer apply. Students may benefit from imagining arithmetic on circular, multidimensional, or other structures. They might then see what kinds of numbers fit these imagined frameworks.

**Author note** Thanks to Kim Au, David Barner, Sybilla Beckmann, Winston Chang, Jacob Dink, Brian Edwards, John Glines, Jacqueline Hirschey, Robert Siegler, Barry Walker, Antonia Yang, Eric Zaslow, and audiences at New York University, the University of Notre Dame, Northwestern University, and the members of a workshop on Semantic Development at the 2010 Cognitive Science meetings for their comments on this research.

**References**


