Developmental Psychology

Development of Fraction Comparison Strategies: A Latent Transition Analysis

Luke F. Rinne, Ai Ye, and Nancy C. Jordan

Online First Publication, February 20, 2017. http://dx.doi.org/10.1037/dev0000275

Developmental Psychology

0012-1649/17/$12.00 http://dx.doi.org/10.1037/dev0000275

Development of Fraction Comparison Strategies: A Latent Transition Analysis

Luke F. Rinne, Ai Ye, and Nancy C. Jordan
University of Delaware

The present study investigated the development of fraction comparison strategies through a longitudinal analysis of children's responses to a fraction comparison task in 4th through 6th grades ($N = 394$). Participants were asked to choose the larger value for 24 fraction pairs blocked by fraction type. Latent class analysis of performance over item blocks showed that most children initially exhibited a "whole number bias," indicating that larger numbers in numerators and denominators produce larger fraction values. However, some children instead chose fractions with *smaller* numerators and denominators, demonstrating a partial understanding that smaller numbers can yield larger fractions. Latent transition analysis showed that most children eventually adopted normative comparison strategies. Children who exhibited a partial understanding by choosing fractions with smaller numbers were more likely to adopt normative comparison strategies earlier than those with larger number biases. Controlling for general math achievement and other cognitive abilities, whole number line estimation accuracy predicted the probability of transitioning to normative comparison strategies. Exploratory factor analyses showed that over time, children appeared to increasingly represent fractions as discrete magnitudes when simpler strategies were unavailable. These results support the integrated theory of numerical development, which posits that an understanding of numbers as magnitudes unifies the process of learning whole numbers and fractions. The findings contrast with conceptual change theories, which propose that children must move from a view of numbers as counting units to a new view that accommodates fractions to overcome whole number bias.

Keywords: development, fractions, magnitude comparison, mathematics learning, strategies

Mastery of fractions is critical for more advanced mathematics learning and achievement (Bailey, Hoard, Nugent, & Geary, 2012; National Mathematics Advisory Panel, 2008; Siegler, Thompson, & Schneider, 2011; Siegler et al., 2012). Unfortunately, fractions are hard for many children (Jordan et al., 2013; Mack, 1995; Mazzocco & Devlin, 2008; Ni & Zhou, 2005; Siegler & Pyke, 2013). To determine why so many children struggle with fractions, it is important to understand how children's knowledge of fractions changes over time as they reach—or fail to reach—proficiency.

A commonly cited source of difficulty with fractions is the overgeneralization of whole number properties, known as "whole number bias" (DeWolf & Vosniadou, 2011; Mack, 1995; Ni & Zhou, 2005; Van Hoof, Verschaffel, & Van Dooren, 2015). For example, judgments of whether 1/3 is greater than 1/2 might be biased by the fact that 3 is greater than 2. Whole number bias not only affects students as they initially learn about fractions (Mack, 1995; Van Hoof et al., 2015), but also interferes with fraction processing for adults who already have relatively well developed fraction understandings (DeWolf & Vosniadou, 2011). When children first learn about fractions, they must recognize, for instance, that unlike whole numbers, larger numbers are often associated with smaller fraction values (i.e., in denominators). For adults with mature fraction knowledge, whole number bias can nonetheless contribute to errors in fraction comparison and computation (DeWolf & Vosniadou, 2015). Such evidence of whole number bias among adults has been cited to support the view that fractions are processed fundamentally as composites of numerators and denominators, rather than numerical magnitudes (DeWolf & Vosniadou, 2015). Whole number bias has also been cited as evidence that children initially possess a core understanding of numbers as *counting units,* an interpretation that does not lend itself to learning of fractions (Ni & Zhou, 2005). Accordingly, a number of researchers have argued that ameliorating whole number bias and understanding fractions and other kinds of rational numbers requires a substantial change of the basic number concept (DeWolf & Vosniadou, 2015; McMullen, Laakkonen, Hannula-Sormunen, & Lehtinen, 2015).

The "conceptual change view" described above has been countered by the integrated theory of numerical development, proposed by Siegler and colleagues (Siegler et al., 2011; Torbeyns, Schneider, Xin, & Siegler, 2015). Proponents of an integrated theory of numerical development argue that while whole number bias may pose some difficulties for early fraction learning, the development of children's number concept does not require extensive conceptual change, because this development is grounded in the recognition that all real numbers—including fractions—can be

Luke F. Rinne, Ai Ye, and Nancy C. Jordan, School of Education, University of Delaware.

This research was supported by Award R305B130012 from the U.S. Department of Education, Institute of Education Sciences.

Correspondence concerning this article should be addressed to Luke F. Rinne, who is now at the Department of Human Development and Quantitative Methods, College of Education, University of Maryland, 3304 Benjamin Building, College Park, MD 20742. E-mail: luke.rinne@gmail.com

represented as numerical magnitudes located on a number line (Siegler et al., 2011). Contrary to the view that fraction processing is fundamentally componential, the integrated theory holds that over time, learners become increasingly able to mentally represent fractions as holistic magnitudes, rather than composites of numerators and denominators (Schneider & Siegler, 2010). Regarding the development of the number concept, Torbeyns et al. (2015) note two distinct ways in which the integrated theory differs from conceptual change theories. First, theories that focus on conceptual change do not recognize that whole number knowledge—in particular, understanding of whole number magnitudes—has a significant positive impact on fraction learning. Second, Torbeyns et al. (2015) point out that overgeneralization of whole number knowledge is not necessarily the only, nor even the predominant difficulty learners face; other difficulties with fraction notation or operations (e.g., multiplication) that cannot be attributed to whole number bias may be even more of an issue.

In this paper, we present the results of a longitudinal study that investigates whether better understanding of whole number magnitudes indeed enhances fractions learning during primary fractions instruction, as suggested by the integrated theory of numerical development. Children completed a commonly used fraction comparison task (i.e., choosing the larger of two fractions; Bailey et al., 2012) in fourth, fifth, and sixth grade; we used latent class and transition analysis to assess patterns and changes in fraction comparison strategies, as well as whether these changes were predicted by whole number line estimation (WNLE) ability after controlling for general math achievement. We also investigate whether children actually do become increasingly able to represent fractions as holistic magnitudes rather than composites of numerators and denominators.

Prior Research on Fraction Comparisons

Although little prior research involving fraction comparison tasks has focused on the issue of conceptual change in fraction learning, many studies have used comparisons to investigate the separate, but related question of how fractions are mentally processed. Bonato, Fabbri, Umilta, and Zorzi (2007) presented adult participants with pairs of fractions with single digit numerators and denominators (e.g., 1/5 vs. 4/5) and examined the magnitudes of distance effects (e.g., Moyer & Landauer, 1967) and effects of spatial numerical association of response codes (SNARC; Dehaene, Bossini, & Giraux, 1993). Bonato and colleagues reasoned that if distance and SNARC effects were more a function of the differences between the whole numbers that constitute numerators and/or denominators—rather than the differences in overall fraction magnitudes—this would provide evidence that fractions tend to be processed as composites rather than magnitudes. Results suggested this was indeed the case, as distance and SNARC effects depended on differences in whole number fraction components (e.g., 1 and 4 in the fractions 1/5 and 4/5), rather than the magnitudes of the fractions themselves.

Schneider and Siegler (2010) and Meert, Grégoire, and Noël (2009, 2010) argue, however, that Bonato et al.'s (2007) results were strongly influenced by their selection of fraction comparison stimuli. The comparisons presented by Bonato et al. were all solvable simply by comparing numerators to numerators (e.g., 2/5 vs. 3/5) or denominators to denominators (1/7 vs. 1/9). Such comparisons do not elicit evaluations of fraction magnitude due to the availability of this very basic strategy—a strategy that would not apply consistently for the broader range of fractions people regularly deal with. Schneider and Siegler (2010) incorporated a variety of fractions with both single digit and double-digit numerators and denominators. They found that adult participants exhibited a logarithmic distance effect, with response times increasing as a function of the arithmetic difference between the values of the fractions in each pair, rather than differences between numerators or differences between denominators. These results indicate that adults can and do represent discrete fraction magnitudes, at least when simpler strategies that focus on numerators or denominators would be ineffective.

Meert et al. (2009) found that comparisons of fractions with identical denominators but different numerators yielded distance effects of whole number numerators, but effects for fraction pairs with identical numerators and different denominators were dependent on differences in fraction magnitude. When fraction pairs did not share any common components, such that direct comparisons of numerators alone or denominators alone would not be sufficient to determine which fraction is greater, representation of fractions as discrete magnitudes increased. Meert and colleagues propose a dual process theory and conclude from their studies that adults likely have "hybrid" representations; that is, they process fraction components individually whenever this is adequate to solve fraction comparisons, but they evaluate fraction magnitudes whenever componential processing is relatively difficult. Obersteiner, Van Dooren, Van Hoof, and Verschaffel (2013) provide additional evidence along these lines showing that expert mathematicians process fraction comparisons componentially and are affected by whole number bias when fractions share numerators or denominators, but when fractions did not share common components, whole number bias was not observed, indicating the use of holistic representations. Obersteiner and Tumpek (2016) supported this conclusion further via an eye-tracking study.

On the other hand, DeWolf and Vosniadou (2015) found that although distance effects were dependent on fraction magnitudes, accuracy and response times for fraction comparisons were nonetheless influenced by consistency with the ordering of whole number fraction components. In addition, they found that as two fractions become closer to one another in magnitude, the influence of whole number orderings increases. DeWolf and Vosniadu argue that as the distance between the two fractions in a given pair shrinks, it becomes more difficult to construct discrete magnitude representations, so access to holistic representations of fractions must therefore be fundamentally indirect, and this indirect access to magnitudes is "lost" when fractions are close to one another.

However, an alternative interpretation of the results found by DeWolf and Vosniadou (2015) is that access to discrete magnitude representations is not *lost* when fractions are separated by only a small difference, but rather, these representations simply become harder to distinguish from one another in the presence of interference from whole number orderings. Recent work indicates that fraction magnitude understanding is supported by a "ratio processing system" (Lewis, Matthews, & Hubbard, 2015; Matthews, Lewis, & Hubbard, 2016) that is analogous to the approximate number system believed to underlie whole number representations (Halberda & Feigenson, 2008; Halberda, Mazzocco, & Feigenson, 2008). Representations of fraction magnitudes are "fuzzy" and overlap as a result of the limited precision of the ratio processing system. Thus, the influence of whole number orderings may increase as fractions become closer together not because discrete holistic representations cannot be constructed, but rather because such representations overlap considerably, reducing the degree of contrast and allowing increased influence from interference effects.

When the research findings described above are considered collectively, it seems reasonable to conclude that during comparison tasks, fractions are in some instances processed as composites of numerators and denominators, and in others as holistic magnitudes. Alibali and Sidney (2015) argue that whether fractions are processed componentially or holistically is a function of a wide variety of factors, including experience with fractions, task demands, context, and the strength and precision of individuals' representations of rational number magnitudes. Accordingly, we believe that for those with a mature understanding of fractions, selective focus on numerators or denominators in isolation may not reflect generalized componential processing, but rather the use of explicit, learned strategies that are appropriate in particular cases (e.g., when denominators are the same and numerators differ). Adoption of new strategies alongside the fluid application of existing strategies is the hallmark of overlapping waves theory (Siegler, 1996)—the view that with development, learners become more adaptive at choosing from a variety of available strategies. Fazio, DeWolf, and Siegler (2016) showed that adults' use of strategies for fraction comparisons is consistent with the overlapping waves theory.

Although the debate over componential versus holistic processing of fractions may revolve around a largely false dichotomy, important questions remain. First, *what mechanisms support children's acquisition* of mature fraction understandings? Researchers who have typically emphasized effects of whole number bias and componential processing (e.g., DeWolf & Vosniadou, 2015) tend to view fractions learning as a matter of wholesale conceptual change. Meanwhile, proponents of the integrated theory of numerical development (e.g., Siegler et al., 2012) view fractions learning as a process of leveraging and expanding on whole number magnitude representations. On this latter view, only isolated, fine-grained conceptual changes are required for fractions learning. However, it has not yet been adequately demonstrated that understanding of whole number magnitudes indeed does facilitate learning of fractions.

The second question remaining is: As children overcome whole number biases and other potential issues with fractions learning, *do they increasingly represent fractions as holistic magnitudes*, at least when simpler strategies are not available? Because studies conducted to date have focused on adults and used cross-sectional designs, results have been interpreted in different ways depending on theoretical viewpoint. What is needed is longitudinal research that investigates changes in children's understanding *during the period when they are initially learning about fractions*. This is the gap in the literature we aim to fill.

The Present Study

Understanding the mechanisms supporting children's development of fraction understanding requires identification of early fractions knowledge and investigation of how this knowledge develops over time. The present study aims to uncover latent trends indicative of strategy use as children solve different types of fraction comparisons. Our fraction comparisons are similar to those used in previous empirical investigations, but we pay particular attention to the different ways in which fraction comparisons are amenable to different strategies. Some comparisons are readily solvable via simple comparisons of single fraction components (e.g., numerators in 3/7 vs. 5/7). However, others do not allow for a simple comparison of components, but rather require the evaluation and comparison of fraction magnitudes. Such evaluations may be conducted via more complex comparisons of fraction components based on sound understanding of how these components relate to magnitude (i.e., simultaneous comparison of numerators to numerators, or denominators to denominators, with recognition of how they function), or through the holistic representation of fraction magnitudes. Importantly, we investigate shifts in strategy among children *who are still in the process of learning fractions*, affording us crucial insights into the cognitive factors that predict the developmental trajectory of fraction understanding.

Our fraction comparison task was administered in fourth, fifth, and sixth grades. Comparisons were blocked at each time point according to the types of fractions being compared (e.g., unit fractions, reciprocals, etc.), such that different strategies (e.g., numerator comparison, denominator comparison, etc.) would be more or less successful across blocks. If participants' accuracy is consistent within blocks, but varies across blocks, patterns of performance across the set of blocks as a whole can reveal more general strategies that children use to determine which fraction they believe is bigger. We used latent class analysis to identify groups of students with distinct response patterns across item blocks, indicating unobserved strategies that are shared within class but differ across classes. We then conducted a latent transition analysis to examine trends in strategy change over time conditional on prior class membership. We tested effects of covariates related to prior mathematics knowledge (general math achievement, whole number line estimation ability), background cognitive factors (e.g., nonverbal reasoning, verbal ability, working memory, etc.), and demographic factors (e.g., gender, SES). Our model includes effects of these covariates not only on class membership, but also on the probability of transitioning among classes over time. Finally, we conducted exploratory factor analyses to investigate whether over time children came to increasingly represent fractions holistically when simpler strategies involving numerator or denominator comparisons were not feasible.

We hypothesized that after controlling for general math achievement and other background variables, whole number line estimation ability measured prior to formal fractions instruction would predict the adoption of normative fraction comparison strategies. Such strategies need not necessarily involve the direct comparison of fraction magnitudes; consistent, appropriate use of strategies that compare numerators or denominators across a range of different fraction types can also be taken as a reflection of normative strategy use and possession of a basic understanding of fraction magnitudes. If whole number line estimation ability has a significant, positive impact on the adoption of normative fraction comparison strategies, this would provide evidence in favor of the integrated theory of numerical development, rather than conceptual change theories, as only the former view provides a reason to believe that whole number magnitude understanding *helps* people build mature fraction understandings. We also predicted that as children grow and acquire a better understanding of fractions, they will be more likely to represent fraction magnitudes holistically to solve fraction comparisons when simpler strategies are unavailable, even if they continue to use simpler comparisons of numerators or denominators when appropriate, consistent with overlapping waves theory (Siegler, 1996).

Method

Participants and Procedure

Study participants were recruited for a broader, 4-year longitudinal study investigating the development of fraction understanding between third and sixth grade (Jordan et al., 2013). Participants came from nine schools serving a socioeconomically diverse student population, and low-income students were oversampled. The fraction comparison task that is the focus of this study was administered beginning in fourth grade (Year 2 of the broader study). Our sample ($N = 431$) includes all students who completed at least the first administration of this task.

The demographics of the sample included in this study were very similar to those for the full group of study participants ($N = 481$). Females comprised 52.2% of the sample, and the average age in third grade was 106.1 months ($SD = 5.5$). The racial composition was 53.2% White, 38.4% Black/African American, 5.5% Asian/Pacific Islander, and 3.0% American Indian/Alaskan Native, with 15.6% additionally identifying as Hispanic. In addition, 10.4% were English language learners (ELL), and 59.7% were classified as low income based on participation in a school lunch program. Beginning in fourth grade, all of the participating schools used mathematics curricula aligned with the Common Core Standards (Common Core State Standards Initiative, 2010).

The fraction comparison task was administered in the fall of fourth grade (T_1), the spring of fifth grade (T_2), and the spring of sixth grade (T_3). The fall of fourth grade serves as a baseline, as few students have received any formal fractions instruction at this point. By the spring of fifth grade, students should have received sufficient instruction to perform all comparisons, according to the standards. Finally, performance in the spring of sixth grade indicates how strategy use may continue to change as fractions are integrated into instruction in other mathematics topics. All background cognitive and demographic measures were administered in third grade.

Measures

Fraction comparison task. The fraction comparison task is a modified and expanded version of the task described by Bailey et al. (2012). Participants were presented with 24 pairs of fractions and asked to circle the larger fraction in each pair. They were instructed to do their best to finish all 24 items within the 3-min time limit. The items were organized into six ordered blocks of four, with all of the items in each block sharing a given type. (See Table 1 for a complete list of items by block.) Participants received a score of "1" if they circled the larger fraction or "0" if they circled the smaller one; the score for each block is the sum over the four items. Nonresponses were treated as incorrect, though these were very rare (< .2% of responses). Given that the time limit allowed children only 7.5 seconds per item, we are confident that this limit prevented children who did not possess a basic understanding of fractions from being able to accurately complete the test within the time limit using labor-intensive rote comparison strategies such as cross-multiplication.

Table 1
Fraction Comparison Task Items, by Item Block

Item block	Item	Fraction pair	
1	1	1/3	1/2
	2	1/55	1/57
	3	1/4	1/5
	4	1/10	1/100
2	5	7/12	9/12
	6	5/7	6/7
	7	24/48	28/48
	8	2/10	4/10
3	9	50/100	16/17
	10	20/40	8/9
	11	5/10	3/4
	12	10/20	5/6
4	13	5/3	5/2
	14	2/4	2/5
	15	6/9	6/12
	16	3/7	3/8
5	17	3/2	2/3
	18	8/4	4/8
	19	5/14	14/5
	20	5/6	6/5
6	21	3/10	2/12
	22	12/50	8/60
	23	6/33	9/30
	24	6/8	3/9

The first block (items 1–4) presented pairs of unit fractions (e.g., 1/3 vs. 1/2). The second block (items 5–8) presented pairs with the same denominator, but different numerators (e.g., 7/12 vs. 9/12). In the third block (items 9–12), one fraction was constructed with relatively large numerators and denominators and equaled one half (e.g., 50/100), whereas the other had smaller numerators and denominators but was just less than 1 (e.g., 16/17). The fourth block included pairs of fractions with the same numerator and different denominators (e.g., 2/4 vs. 2/5), and the fifth block paired reciprocals with one another (e.g., 3/2 vs. 2/3). Finally, the sixth block presented pairs in which numerators and denominators differed in opposite directions (e.g., 3/10 vs. 2/12).

The item blocks were designed to collectively elicit divergent response patterns corresponding to differential strategy use. For instance, if children were simply choosing fractions with larger numbers in both numerators and denominators (a reasonable strategy in the absence of fraction understanding), they would do well on Block 2 (where only numerators differ), but poorly on Blocks 1, 3, and 4. Over the last two blocks, we would expect chance performance; Block 5 presents reciprocals, which have the same numbers across the two fractions (e.g., 6/5 vs. 5/6), while Block 6 presents fractions in which numerators and denominators differ in opposite directions. The use of other kinds of strategies instead or in addition (e.g., focusing on numerators only) would yield different patterns of responses across the set of item blocks. Because different blocks are amenable to different kinds of simple strategies, reaching correct solutions across all item blocks within a limited amount of time demands at least a basic understanding of fraction magnitudes that allows for diverse comparisons of fractions.

Mathematics ability. Additional measures of mathematics ability were included to investigate how the development of fraction comparison strategies relates to mathematics achievement in general and whole number magnitude understanding in particular.

Wide Range Achievement Test. The mathematics subsection of the Wide Range Achievement Test (WRAT; Wilkinson & Robertson, 2006) was used to test children's general ability within the domain of mathematics. The test focuses primarily on computation (addition, subtraction, multiplication, and division) involving whole numbers. Although there are a small number of items that include fractions, very few of these items were attempted by participants in the present study, who had not yet begun formal fractions instruction. Reliability of the WRAT mathematics test is high ($\alpha = .90$), and scores correlate with those of other broad measures of mathematics achievement (Wilkinson & Robertson, 2006).

Whole number line estimation. This task was used to assess children's understanding of whole number magnitudes. After receiving instructions and completing a practice trial, children were presented with 25 cm number lines with end points of 1 and 1000 and were asked to locate a sequence of 22 different numbers (56, 606, 179, 122, 34, 78, 150, 938, 100, 163, 754, 5, 725, 18, 246, 722, 818, 738, 366, 2, 486, and 147). The absolute distance between the location selected by the child and the correct location was averaged across the 22 items, and the percentage of absolute error (PAE) was calculated. As measured by Cronbach's alpha, reliability for this task is .89 (Jordan et al., 2013).

Background variables. Cognitive and demographic variables were included in the analysis to control for incidental effects and to investigate potential relationships to fraction comparison strategies used at each time point and changes in strategies over time.

Nonverbal reasoning. To assess nonverbal reasoning, we used children's age-scaled scores on the Matrix Reasoning subtest of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). In this task, children are presented with a sequence of 2-by-2 grids depicting geometric patterns in three cells. Children are tasked with choosing the next grid in the sequence from five alternatives. The Matrix Reasoning subtest exhibits high reliability ($\alpha = .90$) as well as a high correlation ($r = .87$) with overall scores on the full WASI (Wechsler, 1999).

Verbal ability. Verbal ability was assessed using children's age-scaled scores on the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 2007). Children are presented a series of words, and the test assessor asks the child to choose the corresponding picture from four alternatives. Reliability for the PPVT is high ($\alpha = .96$), and research by Dunn and Dunn (2007) demonstrated strong criterion validity.

Working memory. The Counting Recall subtest of the Working Memory Test Battery for Children (WMTB-C; Pickering & Gathercole, 2001) presents children with a sequence of dot arrays and asks them to remember, in order, the number of dots presented in each array. After children respond correctly to three out of six trials, an additional array is added to the sequence. Test–retest reliability for the Counting Recall subtest is .61 (Pickering & Gathercole, 2001).

Attentive behavior. For each participating child, his or her mathematics teacher completed the SWAN-I Rating Scale (Swanson et al., 2006), which includes nine items designed to assess inattentiveness associated with attention deficit-hyperactivity disorder, based on criteria given in the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (First, 1994). Teachers rate each student on a scale from 1 (*below average*) to 7 (*above average*) for each item. Cronbach's alpha for the SWAN is .97.

Reading fluency. To assess reading fluency, we used children's age-scaled scores on the Sight Word Efficiency subtest of the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999). For this test, students are given 45 seconds to read aloud as many of 104 written words as possible, with the raw score comprising the total number of correctly read words. Test–retest reliability in third grade is .97 (Torgesen et al., 1999).

Demographic variables. We also included children's age in our analysis, as well as binary predictors for gender and income status (measured by participation in a free or reduced cost school lunch program).

Data Analytic Strategy

The purpose of our study was to delineate the strategies children use to compare fractions and to investigate shifts in strategy over time. Our goal was to construct a longitudinal latent variable mixture model and perform a latent transition analysis (LTA). LTA is a longitudinal extension of latent class analysis (LCA), a multivariate statistical approach based on measurement theory in which unobserved, underlying grouping variables (i.e., latent classes) are inferred from a set of categorical indicators (Goodman, 1974; Lazarsfeld, Henry, & Anderson, 1968). Conventional LTA simultaneously defines the latent class structure and models individual-level changes in class membership over time (Collins & Flaherty, 2002).

Preliminary LCA. Prior to conducting our LTA, it was necessary to decide how many latent classes were needed to model the indicator variables, forming the measurement model. To meet the assumptions of LCA, we dichotomized the score for each block of four items in the following way: If a child got two or fewer items in a given block correct, they were assigned a score of 0 for that particular block, indicating that their performance was at or below chance. (Random guesses have a 50% chance of correctness on each item, yielding item-block scores with an expected value of two.) If a child got three or four items correct, they were assigned a score of 1 for that block, indicating better than chance performance.

There were several reasons for dichotomizing the dependent variable as described above and conducting an LCA, rather than a latent profile analysis (LPA; Collins & Lanza, 2010) in which scores from 0 to 4 within each block would be treated as a continuous variable. First, LPA assumes that distributions of continuous indicators within each latent class are normal. However, we expected that scores within each block would frequently be at ceiling (4) or floor (0), as children would likely use the same strategy (e.g., choosing the fraction with the larger numbers) across all four items within a given block. Thus, class-conditional distributions of continuous indicators would likely be highly skewed, undermining the assumptions of LPA. MacCallum, Zhang, Preacher, and Rucker (2002) cite extreme skewness of this sort as a justification for dichotomizing data. In addition, procedures for model selection, evaluation, and the inclusion of covariates in LPA are not well established in the methodological literature, particularly in the context of LTA.

An alternative LCA approach might be to treat each item as an individual indicator; however, each individual response would be a less reliable predictor of class membership than would a binary indicator that aggregates responses across the four items in a given item block. Further, with 24 items data sparseness would become a problem, and there would be too many parameters to estimate. A final alternative approach would be to treat each score above 0 as a categorical threshold; however, this would even more dramatically increase the required number of parameter estimates such that the inclusion in our model of three time points along with covariates would not be possible given our sample size. Thus, dichotomizing responses for each block to indicate above chance versus at/below chance performance was the best option for analyzing the particular data we collected.

We fit a series of LCA models at T_1, T_2, and T_3 to determine the number of classes to be used in the subsequent LTA, relying on a combination of statistical, theoretical, and practical grounds. We first specified a 2-class LCA model and increased the number of classes until models no longer converged or fit the data more poorly. Fit indices were then considered along with interpretations of class structures to choose the number of classes that yielded the most meaningful, parsimonious, and statistically sound model (Collins & Lanza, 2010). The Bayesian Information Criterion (BIC) and Bootstrapped Likelihood Ratio Test (BLRT) have both been recommended as the most reliable measures of fit for determining the appropriate number of latent classes (Nylund, Asparouhov, & Muthén, 2007). Strong latent class separation—manifested in highly distinct latent classes and posterior class membership probabilities near 0 and 1—is also critical for LCA models. The degree of difference among included classes is measured by the entropy statistic, which ranges from 0 to 1, with larger values indicating better latent class separation (Collins & Lanza, 2010). If the measurement model has a low entropy value (e.g., <.6), this suggests that the latent class structure is a poor representation of the dependent variable, and analyses involving latent class membership will be unreliable.

Latent transition analysis. After conducting our preliminary LCAs, we constructed an autoregressive model to estimate *latent transition probabilities*—the probabilities of particular patterns of class change across time points (Collins & Lanza, 2010). Although it is possible to allow for different numbers of latent classes at each time point, doing so can make it difficult to interpret transitions among classes across time points, because the meaning of membership in a particular class may fail to remain the same. Thus, we adopted the same number of latent classes at each time point to simplify interpretations of transitions, as has been previously recommended in LTA research (e.g., Lanza, Patrick, & Maggs, 2010; Nylund, 2007).

Tests of measurement invariance. Prior to estimating latent transition probabilities, it was important to investigate the common assumption of full measurement invariance—that is, the assumption that latent class indicators produce the same classifications across time points. Although we assumed that the number of classes was constant over time, we directly tested the assumption that the parameters of the measurement model remain constant (i.e., conditional item-response probabilities underlying the latent class structure are equivalent regardless of test time). Although this assumption has the benefit of simplicity and increases the degrees of freedom, it may be questionable more often than many researchers realize (Nylund, 2007). A further advantage of relaxing measurement invariance is that doing so allows for the possibility that the factor structure underlying individuals' responses for particular item blocks changes over time as well. Provided that some degree of measurement noninvariance is observed, exploratory factor analyses (EFA) can then be used to investigate whether children increasingly come to represent fractions as holistic magnitudes rather than composites of numerators and denominators.

Rather than lifting the assumption of measurement invariance completely, however, which would lead to a full noninvariant model and the need to estimate a larger number of parameters, we chose to test for *partial* measurement invariance—that is, a scenario in which some parameters of the measurement model are held constant, whereas others are allowed to vary freely (Nylund, 2007). Tests for measurement invariance were conducted prior to including longitudinal relationships between latent variables or covariate effects in the model. We imposed the measurement invariance constraint on each indicator variable one by one, at each stage conducting a likelihood ratio test to see whether adding the constraint significantly reduced the predictive power of the model. After testing measurement invariance for the first transition (i.e., T_1 to T_2), we repeated this procedure for the second transition (i.e., T_2 to T_3).

One downside to allowing measurement invariance is that even if the number of classes is held constant, interpretations of class structure could potentially change if items tend to function differently at different time points. Differential item functioning that could threaten interpretations of class structure would be accompanied by significant changes in conditional response probabilities (CRP)—the probability of a given indicator response conditioned on class membership. Thus, if large changes in CRPs showed that the class structures differ qualitatively across time points, we would then impose measurement invariance on the model regardless of invariance tests to ensure results maintain interpretability.

The three-step LTA model with auxiliary variables. The inclusion of auxiliary variables (e.g., covariates) in LCA and LTA models can be problematic if latent class membership, latent transition probabilities, and effects of covariates on each of these are all estimated simultaneously in a single step (Asparouhov & Muthén, 2014; Vermunt, 2010). The primary issue is that covariate relationships can have a marked influence on the estimation of the measurement model. Our goal was to *observe* a latent class structure, and *then* investigate how covariate values relate to both the latent class structure at particular time points and the probabilities of transitioning between latent classes across time points. Consistent with this research goal, Asparouhov and Muthén (2014) and Nylund-Gibson, Grimm, Quirk, and Furlong (2014) recently detailed a three-step procedure utilizing the structural equation modeling software M*plus*. Simulation studies (Asparouhov & Muthén, 2014; Ye & Rinne, 2016) have shown that this new three-step method works well when the class separation is good (i.e., entropy is 0.6 or higher).

The three-step method for LTA proceeds as follows: First, separate LCAs are conducted at each time point with the same number of classes specified. In the second step, the most likely latent class memberships are derived for each individual from the results of these LCAs (along with error terms). Thus, the measurement model for each time point is established prior to introducing any autoregressive relationships or covariate influences to the analysis. Then, in the third step, an LTA is conducted in which the indicator variables are the most likely class membership values at each time point, rather than the indicator variables used in the initial LCAs. It is only when this final LTA is conducted that covariates and autoregressive relationships are introduced to the model. This ensures that the measurement model is not unduly influenced in the process of constructing a structural model that includes covariates.

We used the three-step LTA procedure to investigate transitions in latent class membership, relationships between covariates and class membership at the initial time point, and relationships of covariates to transition probabilities between time points. Our model includes only first-order autoregressive relationships (i.e., no regressions of T_3 class membership on T_1 class membership), as we see no clear theoretical reason for believing that higher-order relationships exist. Thus, the final three-step LTA model tested in the current study is a three-time-point longitudinal mixture model that estimates transition probabilities from T_1 to T_2 and T_2 to T_3.

The LTA model was tested in M*plus* 7.1 using full information maximum likelihood (FIML) estimation, which maximizes power by assigning values to certain kinds of cells with missing data based on all of the other available data for a given participant. However, FIML estimation in M*plus* is unable to assign values for missing covariates, so data for the LTA were drawn from the 394 participants for whom we had complete covariate information.

Results

Preliminary LCA

We constructed a series of LCA models at each time point to determine the appropriate number of classes to use to model the population. BLRT tests (see Table 2) showed that at T_1, adding a fifth latent class significantly improved model fit, but adding a sixth did not. At T_2 and T_3, BLRTs indicated improved model fits up to four classes. We also examined the BIC; as values decrease, this indicates an increasingly optimal balance between model fit and parsimony (Schwarz, 1978). BIC values suggested more conservative extraction of additional classes, pointing to 4-class, 3-class, and 3-class models for T_1, T_2, and T_3, respectively. The right side of Table 2 gives the counts (proportions) of individuals who fell into each latent class for the model with a given number of classes. Classes 2′ and 3′ are designated as such because they appear to be subclasses of 2 and 3, whose members are uniformly absorbed into those classes as the number of latent classes is reduced. In the 5-class model at T_1, the 7 individuals in Class 2′ join Class 2 if the number of latent classes is reduced to 4. If the number of latent classes is reduced further to three, the 100 individuals in Class 3′ join Class 3.

The similarities and differences among the five classes at T_1 can also be seen in Figure 1, which shows the probabilities of above chance performance by class and item block. Class 2 appears to include individuals who chose the fraction that contains larger numbers, with little to no regard for whether they appeared in the numerator or denominator. Performance was near floor in Blocks 1 and 3, cases in which fraction pairs differ only in the denominator, as well as Block 4, where the smaller fraction has both a larger numerator and a larger denominator. In contrast, performance was near ceiling for Block 2, where only denominators differ. In Block 5 (reciprocals) and Block 6 (numerators and denominators differing in opposite directions), performance was closer to chance, as the strategy of choosing fractions with the larger numbers does not provide a definitive answer. However, the fact that performance remained above chance suggests that members of this class may tend to focus more on the numerator.

Class 2′ differs from Class 2 only in performance on the final two blocks, where Class 2′ was near floor. Thus the difference between Classes 2 and 2′ appears to simply be that members of the former focused more on which number is larger in the numerator, whereas the members of the latter focused on which number is larger in the denominator. Thus, when the number of classes is reduced to three, Class 2′ is absorbed into Class 2 to form a broader "larger number bias" (LNB) class that includes individuals with varying degrees of focus on the numerator versus the denominator.

Class 3, meanwhile, appears to reflect a "smaller number bias" (SNB) combined with a relatively strong bias toward focusing on the numerator. Performance was near ceiling in Blocks 1, 3, and 4,

Table 2
Preliminary LCA Models: Model Fit, Class Separation, and Most Likely Latent Class Membership Counts, by Test Time and Model Type (Extracted vs. 3-Class)

# Classes (*c*), by test time	Model comparisons BLRT $\chi^2(7)$: *c* vs. *c* − 1	*c* vs. *c* + 1	BIC	Ent	Membership count (proportion), by class: 1	2	2′	3	3′
T_1									
5	19.32**	12.26	2354	.91	81 (.19)	203 (.47)	100 (.23)	40 (.09)	7 (.02)
4	92.64**	19.32**	2331	.89	79 (.18)	203 (.47)	100 (.23)	49 (.11)	
3	109.09**	92.64**	2381	.97	80 (.19)	304 (.71)		47 (.11)	
T_2									
4	34.60**	12.32	2154	.88	168 (.46)	66 (.18)	63 (.17)	72 (.20)	
3	150.59**	34.60**	2147	.88	168 (.46)	116 (.31)		85 (.23)	
T_3									
4	19.83**	12.58	1851	.85	183 (.58)	69 (.22)	15 (.05)	51 (.16)	
3	102.99**	19.83**	1830	.87	183 (.58)	78 (.25)		57 (.18)	

Note. LCA = latent class analysis; *T* = Test time (1 = Fall 4th Grade, 2 = Spring 5th Grade; 3 = Spring 6th Grade); *c* = number of classes retained; BLRT = bootstrapped likelihood ratio test; BIC = Bayesian information criterion; Ent = Entropy Norm = normative.
** $p < .001$.

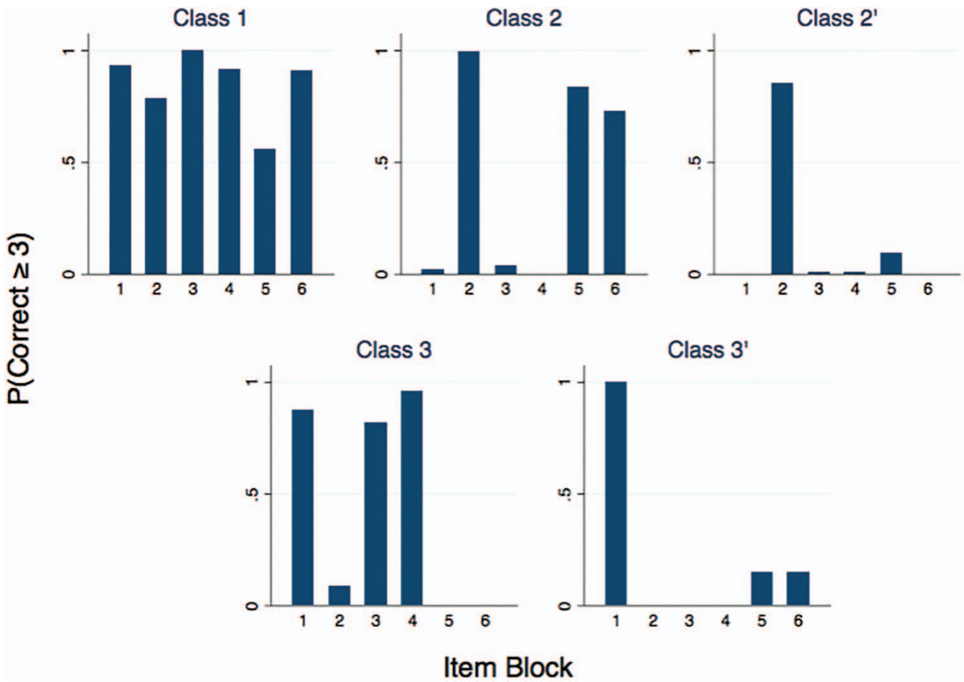

Figure 1. Probabilities of above chance performance at T_1 by class and item block, five-class LCA model. See the online article for the color version of this figure.

where smaller numbers are uniformly associated with larger fractions, and at floor in Blocks 2, 5, and 6, where the larger numerator determines the larger fraction. Class 3′ is somewhat idiosyncratic, with near ceiling performance on the first block, and near floor performance on remaining blocks. Perhaps these individuals were mixing strategies over item blocks. Regardless, this group's near-ceiling performance in Block 1 and near-floor performance in Blocks 2, 5, and 6 leads to absorption into Class 3 when the number of classes is reduced. Although the lack of a clear explanation for this class's performance may seem problematic, keep in mind that only seven of the 341 total participants fell into this class.

Table 2 shows that at T_2 and T_3, when the number of classes is reduced from four to three, individuals in Class 3′ predominantly move to Class 3. Considering that BIC values favor the three-class models at T_2 and T_3, the orderly manner in which the T_1 classes folded into one another led to the conclusion that a three-class model was most appropriate for our latent transition analysis (LTA). Based on the graphs of response probability patterns shown in Figure 2, interpretation of the resulting three-class structure is straightforward. Class 1 includes individuals who exhibited mostly correct comparisons (indicating normative strategy use) across all blocks. Class 2's responses suggested a general strategy of choosing fractions with larger numbers in *either* the numerator *or* denominator, consistent with a broad whole number bias. Because this group included the vast majority of children at T_1, we view this class as the most basic. Finally, children in Class 3 tended to choose the fraction with the *smaller* numbers, suggesting they may

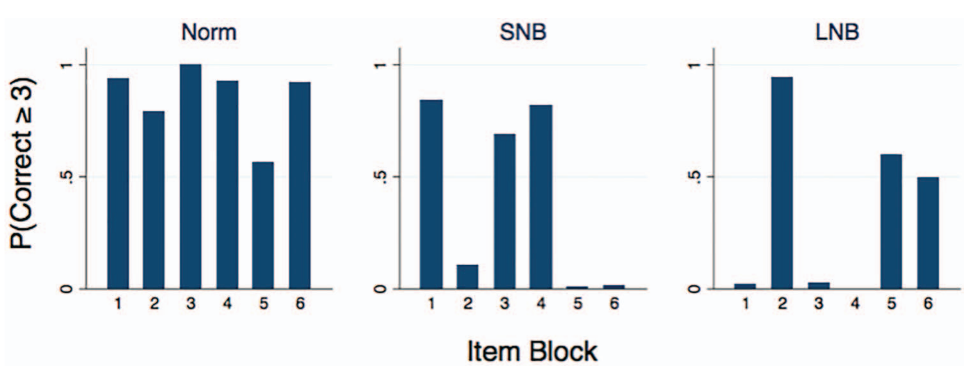

Figure 2. Probabilities of above chance performance at T_1 by class and item block, three-class LCA model. Norm = Normative; SNB = small number bias; LNB = large number bias. See the online article for the color version of this figure.

Table 3
Preliminary LCAs: Predicted Trend in Probability of Above Chance Performance (↑, ↓, or ↔) and Observed Conditional Response Probabilities, by Item Block, Latent Class, and Test Time

		P(# Correct ≥ 3)								
		Norm			SNB			LNB		
Block	Item form	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
1	*1/x* vs. *1/y*	↑ .94	↑ .99	↑ 1.00	↑ .84	↑ .94	↑ .88	↓ .02	↓ .03	↓ .21
2	*x/a* vs. *y/a*	↑ .79	↑ .95	↑ .91	↓ .11	↓ .37	↓ .31	↑ .95	↑ .94	↑ .82
3	*a(1/2)* vs. *(b − 1)/b*	↑ 1.00	↑ .98	↑ .97	↑ .69	↑ .74	↑ .87	↓ .03	↓ .08	↓ .15
4	*a/x* vs. *a/y*	↑ .92	↑ .96	↑ .93	↑ .82	↑ .66	↑ .65	↓ .00	↓ .00	↓ .08
5	*a/b* vs. *b/a*	↑ .56	↑ .90	↑ .95	↔ .01	↔ .24	↔ .12	↔ .60	↔ .56	↔ .49
6	*a/x* vs. *b/y*	↑ .92	↑ .99	↑ .99	↔ .01	↔ .37	↔ .42	↔ .50	↔ .41	↔ .49

Note. LCA = latent class analysis; T = Test Time (1 = Fall 4th Grade, 2 = Spring 5th Grade, 3 = Spring 6th Grade); Norm = normative; SNB = smaller number bias; LNB = larger number bias; ↑ = trend toward correctness predicted; ↓ = trend toward error predicted; ↔ = chance performance predicted.

possess a misconceived partial understanding of fractions; their performance shows some awareness that smaller numbers can somehow lead to larger fraction values (as with denominators), but they did not appear to understand the joint function of the numerator and denominator well enough to consistently give correct responses. Thus, this class may represent a transitional stage.

Table 3 shows expected trends toward correctness or incorrectness based on assumptions of smaller number bias, larger number bias, and normative strategy use. The accompanying observed conditional response probability matches the expected trend direction in nearly every case. Note also that the CRPs in Blocks 5 and 6 (the only ones to defy the expected trend) move closer to chance at T_2 and T_3, indicating that this group's apparent bias toward focusing on the numerator weakened over time, leaving a bias toward smaller numbers as the defining characteristic. Figure 3, which gives the *proportion correct* (out of 4 total) for each class and test time, shows graphically the qualitative similarity of the three-class structure across all three time points. This graph also shows that the qualitative pattern is very similar regardless of whether indicator variables are treated as continuous scores or dichotomized into above chance versus at/below chance performance. Accordingly, the LTA we describe next is conceptualized in terms of a three-class measurement model that changes slightly over time, but which can still be interpreted in terms of the same three broad class types: normative (Norm), smaller number bias (SNB), and larger number bias (LNB).

Three-Step LTA

Structural model constraints. We assumed that children would not regress from the normative class to either the larger number bias class or the smaller number bias class, and thus, the model was constrained such that transition probabilities from Class Norm to Class SNB and Class LNB were fixed to zero. This served to simplify the model both conceptually and computationally.[1]

Measurement invariance. Results for tests of measurement invariance are presented in Table 4. Here, significant results indicate that we should *not* impose the relevant constraint, while nonsignificant results indicate constraints to be appropriate. As shown, imposing measurement invariance from T_1 to T_2 for Blocks 2, 5, and 6 significantly diminished model fit, but this was not the case for Blocks 1, 3, and 4. Meanwhile, measurement invariance held for all blocks except Block 1 for the transition from T_2 to T_3. This suggests that although the three-class measurement structure held over time, the relationships between indicator variable response patterns and class membership changed at least modestly. This can be seen in Table 3 in within-class changes in CRP for individual indicators. However, it is also clear that these changes in the measurement model over time were relatively minor, and the overall interpretation of each class in the three-class structure (Norm, SNB, and LNB) remains consistent, as shown in Figure 3. Thus, we constructed our three-class LTA model on the assumption of a partially invariant measurement structure. Potential sources of partial measurement noninvariance were investigated via exploratory factor analyses reported in a later section.

[1] A model that did not include this constraint yielded a nonpositive definite first order derivative product matrix, a sign of model nonidentification.

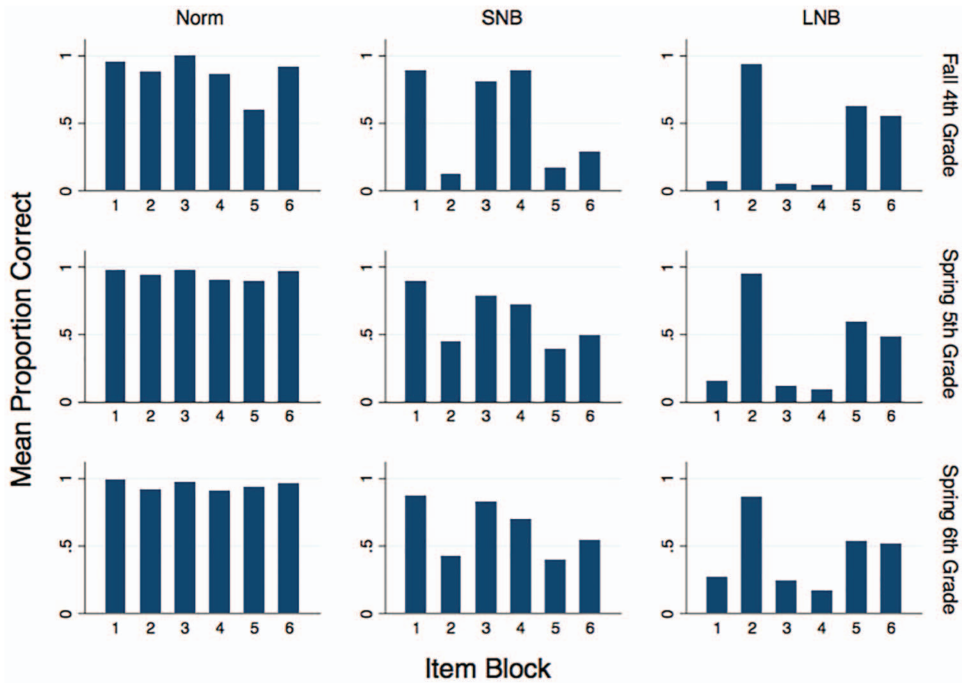

Figure 3. Mean proportion correct by item block, class, and test time. Norm = Normative; SNB = small number bias; LNB = large number bias. See the online article for the color version of this figure.

Class membership and latent transition probabilities. The Step 1 LTA model (which did not include covariates) exhibited a high entropy level of .77 (well above the recommended acceptable level of .6), indicating that classification quality was suitable for including covariates via the three-step LTA procedure. After completing steps 2 and 3 to account for classification error and include covariates, the final model achieved even better classification quality, with an entropy value of .85. The counts of class membership across the three time points (fall of fourth grade, spring of fifth grade, and spring of sixth grade) are given in Table 5. Overall, there was a broad increase in the number of students in the Norm class and a complementary decrease in the number of students in the LNB class. However, this trend does not pinpoint which students moved to which classes at which times; to assess the chances that a given student transitioned from one class to another over a particular time interval, we must examine the latent transition probabilities. Estimated latent transition probabilities from T_1 to T_2 and T_2 to T_3 are given in Table 6.

Table 4
Likelihood Ratio Tests of Measurement Invariance, by Indicator (Item Block) Constraint

Indicator (item block)	T_1 to T_2		T_2 to T_3	
	$\chi^2(6)$	p value	$\chi^2(6)$	p value
1	5.90	.117	12.42	.006**
2	14.26	.003**	6.02	.111
3	5.42	.143	6.92	.074
4	3.53	.317	5.08	.166
5	12.00	.007**	2.91	.406
6	18.65	<.001***	1.83	.608

Note. Significant p values indicate that the assumption of measurement invariance would significantly diminish model fit.
** $p < .01$. *** $p < .001$.

Although a plurality of children in Class LNB at T_1 remained in this class at T_2, a large number transitioned to the Norm class. This represented the group of students who successfully made the transition from the use of a largely naïve comparison strategy to normative strategy use by the end of fifth grade, at which point all common core fractions standards necessary for the comparison of fractions had been covered in school. The number of students reaching the normative class between the spring of fifth grade and spring of sixth grade was smaller, which is to be expected given that at this point the math curriculum has largely moved on to topics other than fractions (although children who struggle with fractions likely continue to receive at least some remedial instruction). Also notice that between the fall of fourth grade and the spring of fifth grade, the number of students in the smaller number bias group greatly increased, but then decreased slightly by the spring of sixth grade. This interesting result suggests that the SNB class appears for many students to have been a stage of partial understanding that was transitional in nature, as children in this class had a higher probability of subsequently reaching the Norm class than did children in the LNB class. However, there were still a relatively large number of students in the SNB class at the end of sixth grade, which indicates that a sizable proportion may have become "stuck" with this strategy, failing to progress further. Finally, there was an even bigger group of students who remained in the LNB class at the end of sixth grade. This finding indicates that there were many students who, despite extensive fractions instruction, never moved beyond the "naïve" LNB strategy.

Table 5
Counts of Most Likely Latent Class Membership by Time

Time	Class		
	Norm	SNB	LNB
Fall 4th Grade	63 (17.1%)	43 (11.7%)	263 (71.2%)
Spring 5th Grade	157 (42.5%)	96 (26.0%)	116 (31.4%)
Spring 6th Grade	197 (53.4%)	87 (23.6%)	85 (23.0%)

Note. Norm = Normative; SNB = smaller number bias; LNB = larger number bias.

When we examine transition probabilities for students who were in Class SNB at either T_1 or T_2, we see that there was only a small probability that they would transition to Class LNB. These findings provide further evidence that the smaller number bias class is often a transitional stage, although it should be noted that students in Class 2 sometimes *did* backslide to Class 3, particularly between T_2 and T_3. A final interesting trend is that there was a much higher probability of transitioning to the Norm class from either the SNB or LNB classes between T_1 and T_2 than between T_2 and T_3. Although this makes sense given that the former represents a longer period of time during which the bulk of instruction required for fraction comparisons occurs, these results may also suggest that as time goes by, the chances of such students ever reaching the Norm class may be quite low in the absence of significant further intervention.

Effects of covariates. Because of limitations of computation and statistical power, it was not possible to include all of the background variables in our LTA model. To determine which to include, we performed a three-step LCA at T_1 and included only those background variables that significantly predicted class membership in at least one pairwise multinomial logistic comparison of classes (see Appendix A for results). Significant effects were found for three background variables: nonverbal reasoning, attentive behavior, and verbal ability. There were no significant effects of gender, age, reading fluency, working memory, or income status. The relations between covariates and class membership at the first time point are given in Table 7 as logistic regression coefficients, with Class LNB serving as the reference group. Odds ratios are also provided to aid interpretation. A correlation matrix for all predictors included in the model is given in Appendix B.

We chose class LNB as the reference group because this is the class in which the majority of students start, and a majority of children transitioned from this class to the SNB or Norm classes at some point between T_1 and T_3, with very few "backsliding" from the SNB class to the LNB class. In addition, it was not feasible to treat the Norm class as the reference class, because the probabilities of transitioning from the Norm class to other classes were fixed to zero to account for the fact that children would almost certainly stay in the Norm class once it was reached. Effects of covariates on transition probabilities cannot be reliably measured in terms of their effects on membership in the normative class, because covariates had no effect on the probability of staying in the normative class, as this probability was fixed to 1.

Note that overall, the positive effects of cognitive covariates tended to be stronger for Class Norm (vs. LNB) than for Class SNB (vs. LNB). This finding is likely related to the fact that Class SNB reflects only a partial understanding of fraction magnitudes, whereas reaching the Norm class requires a deeper understanding, drawing a sharper contrast with the LNB class. Of particular interest is the presence of effects on initial class membership of Math Achievement and the lack of effects of WNLE accuracy. This suggests that general math achievement is highly predictive of the comparison strategies children start with, while more specific skills related to whole number knowledge are not predictive.

The coefficients in Table 8 indicate how the covariates in question *alter* the odds given in Table 7 at each subsequent time point by influencing the likelihood of transitioning from one class to another. Of particular importance here, the results show that WNLE was a significant predictor of transitioning from the LNB class to the Norm class across the first two time points, whereas Math Achievement was not a significant predictor of this transition. Each 1-point increase in WNLE PAE was associated with a 12% decrease in the odds of transitioning from the LNB to the Norm class. Given that the standard deviation of WNLE PAE was 6.7, a one standard deviation reduction in PAE makes a child 2.14 times more likely to transition from the LNB class to the normative class. Thus, although Math Achievement was the stronger predictor of what strategy children *started* with, WNLE ability was the more important predictor of *change* from the more basic LNB strategy class to the Norm class. An analogous effect of WNLE on transition probability was not observed over the second interval between T_2 and T_3. However, we note that while 46% (83) of the students changing classes between T_1 and T_2 changed from the LNB class to the Norm class, this was true of only 14% (12) of the students who changed classes between T_2 and T_3. Thus, the effect of WNLE across the first two time points reflects the vast majority of the students who transitioned from LNB to Norm overall, and the number of students who fell into this category between T_2 and T_3 is too small to draw meaningful conclusions.

In general, cognitive covariates were positively associated with transitions from Class LNB to either the Norm or SNB classes. In particular, attentive behavior and verbal ability, along with WNLE, appear to be closely related to making the direct transition between the LNB class and the Norm class. Surprisingly, for children who were in Class SNB at T_2, coefficients for nonverbal reasoning were significantly *negative* with respect to the chances of transitioning to the normative class or staying in the smaller number bias class. It should be noted, however, that the group of students who transitioned to LNB from SNB across the second interval was relatively small (16), so it is quite possible that these children simply had anomalously high Nonverbal Reasoning scores.[2]

Exploratory Factor Analysis

To investigate potential sources of measurement noninvariance, we conducted a series of exploratory factor analyses aimed at understanding how the factor structures underlying fraction comparison performance changed over time. Factor analyses were conducted on the full set of individual items with the goal of identifying factors or sets of factors that reflect the use of particular strategies at each time point. As suggested by Thompson and Daniel (1996), we considered multiple criteria to determine the

[2] An alternative explanation is that greater nonverbal reasoning makes children less likely to maintain a partial misconception (smaller number bias) and/or more likely to revert to their naïve conception (larger number bias).

Table 6
Estimated Latent Transition Probabilities by Time

Transition period	Starting class	Ending class		
		Norm	SNB	LNB
T_1 to T_2	Norm	1.000	.000	.000
	SNB	.460	.489	.052
	LNB	.270	.304	.426
T_2 to T_3	Norm	1.000	.000	.000
	SNB	.288	.566	.146
	LNB	.178	.280	.542

Note. The probability of staying in Class Norm is constrained to be equal to 1, so the probability of transitioning from Norm to SNB or LNB is 0. Norm = Normative; SNB = smaller number bias; LNB = larger number bias.

proper number of factors at each time point, including statistical data (i.e., factor eigenvalues, cumulative proportions of explained variance), as well as graphical representations (scree plots; Cattell, 1966) and interpretability. This process yielded three-factor structures for T_1 and T_2, and a four-factor structure for T_3. Eigenvalues and variance proportions for each retained factor are given in Table 9. Oblique oblimin ($\gamma = 0$; Jennrich, 1979) rotated factor loadings are presented in Table 10. For information about specific fraction comparison items, refer back to Table 1.

As shown in Table 10, the factor loadings at all time points corresponded very clearly to particular sets of item blocks, further validating the blocked structure of the test. Blocks 1, 3, and 4 loaded exclusively on Factor 1, which therefore likely reflects successful use of a "smaller numbers" strategy. Meanwhile, Block 2 loaded exclusively on Factor 2, corresponding to the successful application of a "larger numbers" strategy. Block 5 loads exclusively on Factor 3, identifying this factor most closely with the solution of comparisons involving reciprocals. Finally, note that Block 6 loads relatively evenly across all three factors, though the loadings are somewhat lower. This results from the fact that each of the simple strategies that may lead to success on more constrained blocks of items (e.g., items in which only numerators or denominators differ) are only sometimes successful when items differ in terms of both numerators and denominators.

As we move to T_2 we see that while three factors were retained, as at T_1, their order and structure changes somewhat; this helps explain observed measurement noninvariance across this interval in the LTA model. At T_2, the factor onto which the reciprocal items load now has the second-greatest eigenvalue, and it is important to notice as well that the items in Block 6 no longer load on this factor. Instead, the items in Block 6 load only on the factors reflecting the use of "smaller numbers" and "larger numbers" strategies. This may indicate that strategies for dealing with reciprocals are becoming more refined (likely because students have now learned how reciprocals relate to one another). One possibility is that instead of using a "larger numbers" or "smaller numbers" strategy that differentially focuses on the numerator or denominator, depending on the individual, children were now viewing these fractions more holistically and recognizing that one of them is greater than one while the other is smaller. In addition, the fact that Block 6 still loads on the factors underlying the "larger numbers" and "smaller numbers" strategies suggests that even those children who are getting the items in Block 6 consistently correct (i.e., the Norm class), may be reaching their solutions by comparing numerators to numerators (choosing the larger) and/or denominators to denominators (choosing the smaller).

Finally, however, at T_3, we see a fourth factor emerge that corresponds specifically to the items in Block 6, in which numerators and denominators differ in opposite directions. Even though the correct choice for all of these items (because of the design of the block) corresponds to the fraction with the larger numerator (or smaller denominator), the fact that these items load uniquely on a completely separate factor indicates that participants generally were not using either a "larger numbers" or "smaller numbers" strategy—nor the strategy they use for reciprocals—to reach the correct solutions for these comparisons. Thus, this factor likely reflects a strategy of representing each fraction holistically and comparing their magnitudes to find the solution.

Discussion

The present study is the first, to our knowledge, to examine how strategies for fraction comparison develop over time between fourth and sixth grades. Proponents of the integrated theory of numerical

Table 7
Effects of Covariates on LTA Model T_1 Class Membership

Class	Covariate	Coefficient	*SE*	*p* value	Odds ratio
Norm	Attentive behavior	−.010	.020	.043*	.99
	Verbal ability	.037	.015	.013*	1.04
	Nonverbal reasoning	.227	.110	.038*	1.25
	Math achievement	.103	.018	<.001***	1.11
	WNLE	−.062	.059	.606	.94
SNB	Attentive behavior	.042	.021	.043*	1.04
	Verbal ability	.026	.015	.073	1.03
	Nonverbal reasoning	−.084	.068	.217	.92
	Math achievement	.048	.021	.022*	1.05
	WNLE	.014	.018	.664	1.01

Note. No results are given for Class LNB (larger number bias), as this was the reference group in the analysis. Given that Class LNB appears to reflect the most basic level of fraction understanding, effects of covariates on membership in Classes Norm and SNB can be interpreted as effects on the chances of being in each of these more "advanced" classes relative to the more basic LNB class. LTA = latent transition analysis; Norm = Normative; SNB = smaller number bias; WNLE = whole number line estimation.
$^{*}p < .05$. $^{***}p < .001$.

Table 8
Effects of Covariates on Latent Transition Probabilities

Transition interval	Starting class	Ending class	Covariate	Est.	*SE*	*p* value	Odds ratio
T_1 to T_2	SNB	Norm (vs. LNB)	N/A	N/A	N/A	N/A	N/A
		SNB (vs. LNB)	N/A	N/A	N/A	N/A	N/A
	LNB	Norm (vs. LNB)	Attentive behavior	.071	.027	.021*	1.07
			Verbal ability	.059	.019	.002**	1.06
			Nonverbal reasoning	.078	.104	.451	1.08
			Math achievement	−.003	.027	.906	1.00
			WNLE	−.126	.054	.019*	.88
		SNB (vs. LNB)	Attentive behavior	.013	.022	.514	1.01
			Verbal ability	−.008	.015	.605	.99
			Nonverbal reasoning	.132	.076	.080	1.14
			Math achievement	.014	.021	.514	1.01
			WNLE	−.007	.029	.813	.99
T_2 to T_3	SNB	Norm (vs. LNB)	Attentive behavior	.044	.052	.399	1.04
			Verbal ability	.092	.049	.059	1.10
			Nonverbal reasoning	−.630	.424	.138	.53
			Math achievement	−.025	.092	.782	.98
			WNLE	−.132	.123	.283	.88
		SNB (vs. LNB)	Attentive behavior	.006	.045	.899	1.01
			Verbal ability	.078	.070	.263	1.08
			Nonverbal reasoning	−.704	.271	.009**	.49
			Math achievement	.029	.056	.610	1.03
			WNLE	.072	.085	.400	1.07
	LNB	Norm (vs. LNB)	Attentive behavior	.157	.160	.324	1.17
			Verbal ability	−.015	.095	.877	.99
			Nonverbal reasoning	.043	.268	.873	1.04
			Math achievement	.113	.088	.202	1.12
			WNLE	−.004	.193	.984	1.00
		SNB (vs. LNB)	Attentive behavior	.025	.062	.683	1.03
			Verbal ability	.041	.076	.588	1.04
			Nonverbal reasoning	−.065	.161	.685	.94
			Math achievement	−.070	.046	.124	.93
			WNLE	−.038	.072	.592	.96

Note. Norm = Normative; SNB = smaller number bias; LNB = larger number bias; WNLE = whole number line estimation; N/A = not available. Because very few participants who started in Class SNB at T_1 transitioned to other classes at T_2, estimates could not be obtained for covariate effects on transition probabilities. Class Norm is not included as a starting class, because the transition probability was fixed to 1 for starting class Norm.
* $p < .05$. ** $p < .01$.

development argue that children learn to represent number in a way that integrates fractions (among other types of numbers) as discrete values within the ordered set of rational (and eventually real) numbers. Previous work has shown that children's ability to represent fraction magnitudes—a critical skill according to the integrated theory of numerical development—predicts both fraction arithmetic proficiency and overall mathematics achievement (Siegler et al., 2011). However, researchers have not yet provided empirical evidence demonstrating that acquisition of fraction magnitude understanding necessary to perform a wide range of comparisons is supported by whole number magnitude understanding.

Our study also investigated how prior mathematical knowledge and cognitive factors relate to the adoption of normative strategies, as well as how new comparison strategies involving holistic representation of fraction magnitudes may emerge over time during the course of fractions instruction. According to overlapping waves theory (Siegler, 1996), children add new strategies alongside existing strategies, shifting toward more adaptive strategies over time. Although Fazio et al. (2016) showed that adults' use of fraction comparison strategies is consistent with overlapping waves theory, that study did not longitudinally investigate the fraction learning *process*. We sought to determine whether the overlapping waves view also describes *children* as they learn fractions in school. This is important, because overlapping waves theory provides a useful explanation for why children may continue to solve some fraction comparisons via componential processing even as strategies relying on holistic representation of fractions become available.

Our statistical methodology enabled us to investigate a variety of phenomena in a highly nuanced manner. Latent class and transition analyses allowed us to delineate the use of broad strategies over different kinds of fraction pairs, revealing children's underlying understanding of fraction magnitudes. We were able to show the strategies children start with, their shifts among these strategies over time, and effects of covariates on strategy use. Loosening assumptions to allow partial noninvariance in our latent transition model allowed not only for the possibility that children may apply general strategies to different sets of items at different points in development, but also invited investigation of potential changes in factor structure indicative of holistic representation of fractions.

The three-step procedure for analysis of covariate effects revealed that general cognitive predictors, including nonverbal reasoning, verbal ability, and attentive behavior, are related both to children's initial strategy (i.e., latent class in the fall of fourth grade) and to their probabilities of transitioning to more advanced strategies (and ultimately the normative class). Our analysis also showed that children's general math achievement is perhaps the best predictor of fraction comparison strategies prior to formal instruction. However, whole number line estimation ability predicted transitions to normative strategy use across the first two time points—even while controlling for effects of other cognitive factors and general mathematics achieve-

Table 9
Eigenvalues and Variance Proportions for Factor Analyses at T_1, T_2, and T_3

	T_1		T_2		T_3	
Factor	Eigenvalue	Variance proportion	Eigenvalue	Variance proportion	Eigenvalue	Variance proportion
Factor 1	13.86	.58	12.05	.50	12.26	.51
Factor 2	6.31	.26	6.78	.28	5.35	.22
Factor 3	1.56	.06	1.35	.06	1.26	.05
Factor 4	—	—	—	—	1.14	.05

ment. A one standard deviation reduction in percent absolute error more than doubled a child's odds of transitioning from the LNB class to the Norm class, lending credence to the view that whole number magnitude knowledge *directly supports* the acquisition of normative fraction comparison strategies. This result is not explained by the theory that overcoming whole number bias requires substantial conceptual change from a conception of whole numbers as counting units to a new conception of number that accommodates fractions.

Our results are best explained by the view that fraction understanding is constructed primarily by leveraging and expanding on the foundation of whole number magnitude understanding. Early understanding of whole number magnitudes may help children reason about fractions in multiple ways. Knowledge of equal interval spacing of whole numbers could provide a basis for learning how whole intervals are subdivided into fractional intervals. In addition, visualizing ratios of whole number values on a number line could also support understanding of fraction magnitudes, consistent with research on the ratio processing system (Matthews et al., 2016).

Table 10
Rotated (Direct oblimin, $\gamma = 0$) Factor Loadings Greater Than .3 at T_1, T_2, and T_3

		T_1			T_2			T_3			
Block	Item	F1	F2	F3	F1	F2	F3	F1	F2	F3	F4
1	1	.92			.94			.87			
	2	.91			.88			.85			
	3	.95			.95			.77			
	4	.97			.65		.56	.52			
2	5		.81				.81				.89
	6		.80				.76				.92
	7		.75				.72				.81
	8		.75				.71				.93
3	9	.96			.92			.91			
	10	.99			.94			.88			
	11	.98			.95			1.00			
	12	.97			.92			.95			
4	13	.78			.75			.73			
	14	.95			.96			.89			
	15	.97			.91			.69			
	16	.96			.88			.70			
5	17			.96		.93			.87		
	18			.93		.92			.80		
	19			.91		.91			.80		
	20			.92		.86			.82		
6	21	.67	.52	.38	.69		.47			.61	
	22	.44	.56	.40	.53		.54			.70	
	23	.43	.56	.43	.57		.63			.86	
	24	.33	.62	.44	.43		.61			.66	

The results of this study provide a detailed depiction of the developmental trajectory of children's fraction comparison strategies. When children begin fourth grade, most appear to possess a relatively naïve view of fractions, one that is, for the time being at least, biased by knowledge of whole numbers; children tend to believe that whichever fraction has larger values in numerators and/or denominators has the larger magnitude (e.g., DeWolf & Vosniadou, 2011; Mack, 1995; Ni & Zhou, 2005; Stafylidou & Vosniadou, 2004). In contrast, a relatively smaller group of students exhibits a bias toward choosing fractions with *smaller* numerators and/or denominators, which suggests a partial understanding of fraction magnitudes. Such students seem to recognize that larger numbers can somehow lead to smaller fraction magnitudes, but they do not fully understand the relationship between the numerator and denominator. For example, they might correctly judge that 1/4 is greater than 1/5 (possibly because of initial instruction with unit fractions), but would mistakenly believe that 5/7 is greater than 6/7 because of 5 being less than 6. Finally, a third group of students appears to have fully normative comparison strategies, with the exception of fractions greater than one. Such knowledge may derive from early exposure to fractional numerical magnitudes prior to the beginning of primary formal fractions instruction in fourth and fifth grade. Future research should investigate the impact of early experiences with numerical information on children's understanding of fraction magnitudes.

As children receive fractions instruction, the influence of whole number bias decreases, a trend reflected in shifts from the larger number bias class to the normative class or the smaller number bias class. The smaller number bias class may represent a stepping stone to normative strategy use; children in this class in the fall of fourth grade were more likely to transition to the normative class by the spring of fifth grade than were those who initially possessed the larger number bias. The novel finding of this kind of "transitional" stage provides additional evidence that the shift from dealing with whole numbers to dealing with fractions may be more incremental in nature than has been suggested by proponents of a strong conceptual change view (e.g., DeWolf & Vosniadou, 2015; McMullen et al., 2015).

Between the spring of fifth grade and the spring of sixth grade, children continued to transition to the normative class, albeit with a considerably smaller probability compared with the previous period. Effects of covariates also largely disappear, which is likely attributable in large part to the fact that transitions among classes became considerably rarer. Although such results may reflect the shorter length of the latter interval (1 year vs. ~1.5 years), they may also suggest that students who have not reached the normative class by the spring of fifth grade are likely to be those students who persistently struggle with fractions. This is consistent with previous work (Mazzocco & Devlin, 2008; Siegler & Pyke, 2013) showing that many low-achieving students make little additional progress in fraction knowledge between sixth and eighth grade.

At the final time point, the spring of sixth grade, only 55% of the children in our sample had acquired the normative fraction comparison strategies one would expect for students with a mature understanding of fraction magnitudes. Meanwhile, about 20% of students were using the smaller number bias strategy, and 25% still exhibited the larger number bias strategy *typical of students who have received no instruction in fractions at all*. Further, because the likelihood of transitioning to the normative class appears to decrease over time, there is little reason to believe that a majority of these students eventually acquire normative comparison strategies of fractions without intervention. Although a small minority of students may always struggle with fractions because they have core mathematics learning disabilities, established criteria for students' mathematics achievement classify those above the 25th percentile as "typical," whereas only those in the 10th percentile or below are considered likely to have a mathematics learning disability (Mazzocco & Devlin, 2008). Thus, there is clearly a large population of students whose mathematics achievement would be considered low or even "typical," but who have not acquired normative fraction comparison strategies within the expected time frame. Only 44% of the students in our study possessed an understanding of fractions that allowed them to consistently make correct comparisons by the end of fifth grade, at which point all of the standards necessary to support this understanding should have been covered (Common Core State Standards Initiative, 2010). Given that SES did not have a significant effect on class membership, it is highly unlikely that this result is solely attributable to our oversampling of low-income students for our study. Thus, although our results indicate that understanding of whole number magnitudes underpins the development of normative fraction comparison strategies, this by no means entails that whole number bias is unrelated to persistent difficulties with fractions. A likely explanation for the continued struggles of such a large proportion of children may be the failure of mathematics instruction to establish sufficient whole number magnitude understanding.

Results from our exploratory factor analyses further show how children's understandings of different kinds of fractions change over time. At the beginning of fourth grade, prior to formal fractions instruction, children tend to view all proper fractions in more or less the same manner, as they appear to apply a "larger numbers" or "smaller numbers" strategy across the board. Improper fractions may be an exception, as evidenced by the presence of a factor associated primarily with comparisons of reciprocals. The notion that improper fractions and/or reciprocal relationships are special appears to solidify by the end of fifth grade, as these items become differentiated in the factor structure from items that differ in both numerators and denominators, which likewise do not lend themselves to a simple "larger numbers" or "smaller numbers" strategy. Finally, by the end of sixth grade, we observe the emergence of a new factor that corresponds uniquely to items that differ in terms of both numerators and denominators. As noted by Meert et al. (2009, 2010) and Obersteiner et al. (2013), if individuals are capable of constructing holistic representations of fraction magnitudes, this behavior will likely be exhibited for comparisons of fractions with different components, as componential approaches are more difficult in such cases. Thus, we see in our results that, at least for the children included in our sample, the capacity to construct holistic representations may tend to emerge sometime between the end of fifth grade and the end of sixth grade.

Several limitations of the present study should be noted. Our study is restricted to students in a single geographic region. Although the sample was quite diverse and likely reasonably representative of American students in general, we did oversample low-income children, and our study population may possess additional idiosyncratic characteristics of which we are unaware. Moreover, we did not assess or observe instructional methods, which may vary across teachers and schools, and we were unable to document precisely when fractions instruction stopped and whether some children continued to receive remedial help. Another limitation is our use of a paper and pencil fraction comparisons test with simple correct/incorrect coding of responses. The pattern of responses across item blocks on this test only captured basic understanding of fraction magnitudes, and observed effects may not generalize to other aspects of fractions learning, such as computational procedures. Blocking of items may also have influenced children's responses by encouraging them to use the same strategy for all items in a given block when this might otherwise not have been the case. Finally, it is possible that children could have solved some comparisons through rote procedures associated with particular fraction types or by more general strategies such as cross multiplication. However, it is highly implausible that children in the Norm class, who performed near ceiling across all item blocks, could achieve this level of performance without a solid understanding of fraction magnitudes, and a general strategy of cross-multiplication could not feasibly be executed for a large number of comparisons given the time limit.

With respect to statistical methodology, we emphasize that assumptions made during construction of our latent class and transition models are well grounded in the mixture modeling literature, but it is possible that different assumptions could potentially alter conclusions. Another limitation derives from the fact that transitions of any kind were relatively rare across the second interval from the spring of fifth grade to the spring of sixth grade, preventing us from drawing conclusions regarding effects of covariates over this period. Finally, as with any exploratory factor analysis, our interpretations of the factor structure are at least somewhat speculative and should be taken with caution.

The results of our study suggest that current U.S. instruction in fractions is inadequate for many students, and emphasis of core ideas about fractions should continue beyond the end of fifth grade, likely being revisited frequently throughout middle school. It also might be helpful for teachers to distinguish different types of broad misconceptions (e.g., larger number bias vs. smaller number bias) and respond accordingly. Whereas students with the larger number bias misconception may need special help just to understand that whole number magnitudes can potentially be inversely related to fraction magnitudes (i.e., when they appear in denominators), students with the smaller number bias misconception simply need to understand that this relationship *only* holds for the denominator, and *not* for the numerator. It is important for teachers to recognize that this may be a relatively common stop on the way to a mature understanding of fractions—therefore, efforts to uniformly discourage smaller number bias may be counterproductive. And throughout the instructional process, teachers should work to make sure whole number properties and fraction properties are not only contrasted, but also explicitly linked to one another in accordance with the finding that understanding of whole number magnitudes *supports*, rather than detracts from the adoption of normative fraction comparison strategies.

It is also quite plausible that educators could help foster the developmental outcomes observed between fourth and sixth grade in our study—in particular the ability to represent fractions holistically—not only by having students practice representing fractions as values on number lines, but also by engaging them in speeded practice of comparisons or other tasks that build their ability to spontaneously see fractions as discrete magnitudes. The ability to represent fractions as magnitudes may open the door to further strategies (such as fraction reduction or approximation), which will help students as they move on to higher-level mathematics courses in which it is assumed that they are fluent with fractions. We also note that the ability to compare reciprocals appears to have a unique place in children's understanding of fraction magnitudes; it appears to emerge before children engage in holistic representation of fraction magnitudes for a broader group of fractions with differing values across numerators and denominators. Reciprocals and improper fractions could support holistic representation of magnitudes, because they lend themselves naturally to comparisons with 1. Currently, early fractions instruction focuses on proper fractions, yet improper fractions have been shown previously to be very difficult for students to comprehend, with many believing their value to be less than one (Resnick et al., 2016).

A final contribution of our study is methodological in nature. We believe that our use of the recently developed three-step LTA procedure (Asparouhov & Muthén, 2014; Nylund-Gibson et al., 2014) is among the first applications of this advanced modeling technique to address the conundrum that arises when auxiliary variables relate to both latent class membership and transitions over time: How can researchers investigate covariate effects on transitions without influencing the formation of the latent class structure, potentially altering its interpretation and threatening validity? Our tests of measurement invariance and subsequent explorations of noninvariance using exploratory factor analysis are likewise important, because they account for and explain nuanced changes in the measurement model. Although the three-step procedure for LTA is analytically complex and computationally challenging, we believe that the statistical modeling techniques presented in this paper can and should be used in a wider variety of settings in developmental research. We hope that our study both motivates future applications of the three-step LTA procedure and serves as a model for such work.

References

Alibali, M. W., & Sidney, P. G. (2015). Variability in the natural number bias: Who, when, how, and why. *Learning and Instruction, 37,* 56–61. http://dx.doi.org/10.1016/j.learninstruc.2015.01.003

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using M plus. *Structural Equation Modeling, 21,* 329–341. http://dx.doi.org/10.1080/10705511.2014.915181

Bailey, D. H., Hoard, M. K., Nugent, L., & Geary, D. C. (2012). Competence with fractions predicts gains in mathematics achievement. *Journal of Experimental Child Psychology, 113,* 447–455. http://dx.doi.org/10.1016/j.jecp.2012.06.004

Bonato, M., Fabbri, S., Umiltà, C., & Zorzi, M. (2007). The mental representation of numerical fractions: Real or integer? *Journal of Experimental Psychology: Human Perception and Performance, 33,* 1410–1419. http://dx.doi.org/10.1037/0096-1523.33.6.1410

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1,* 245–276. http://dx.doi.org/10.1207/s15327906mbr0102_10

Collins, L. M., & Flaherty, B. P. (2002). Latent class models for longitudinal data. *Applied Latent Class Analysis, 28,* 7–303.

Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis*. Hoboken, NJ: Wiley.

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

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122,* 371–396. http://dx.doi.org/10.1037/0096-3445.122.3.371

DeWolf, M., & Vosniadou, S. (2011). The whole number bias in fraction magnitude comparisons with adults. In L. Carlson, C. Hoelscher, and T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1751–1756). Austin, TX: Cognitive Science Society.

DeWolf, M., & Vosniadou, S. (2015). The representation of fraction magnitudes and the whole number bias reconsidered. *Learning and Instruction, 37,* 39–49. http://dx.doi.org/10.1016/j.learninstruc.2014.07.002

Dunn, D. M., & Dunn, L. M. (2007). *Peabody picture vocabulary test: Manual* (4th ed.). Minneapolis, MN: Pearson.

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

First, M. B. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Goodman, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika, 61,* 215–231. http://dx.doi.org/10.1093/biomet/61.2.215

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "Number Sense": The Approximate Number System in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology, 44,* 1457–1465. http://dx.doi.org/10.1037/a0012682

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

Jennrich, R. I. (1979). Admissible values of γ in direct oblimin rotation. *Psychometrika, 44,* 173–177. http://dx.doi.org/10.1007/BF02293969

Jordan, N. C., Hansen, N., Fuchs, L. S., Siegler, R. S., Gersten, R., & Micklos, D. (2013). Developmental predictors of fraction concepts and procedures. *Journal of Experimental Child Psychology, 116,* 45–58. http://dx.doi.org/10.1016/j.jecp.2013.02.001

Lanza, S. T., Patrick, M. E., & Maggs, J. L. (2010). Latent transition analysis: Benefits of a latent variable approach to modeling transitions in substance use. *Journal of Drug Issues, 40,* 93–120. http://dx.doi.org/10.1177/002204261004000106

Lazarsfeld, P. F., Henry, N. W., & Anderson, T. W. (1968). *Latent structure analysis*. Boston, MA: Houghton Mifflin.

Lewis, M. R., Matthews, P. G., & Hubbard, E. M. (2015). Neurocognitive architectures and the nonsymbolic foundations of fractions understanding. In D. Berch, D. Geary, & K. Mann-Koepke (Eds.), *Development of mathematical cognition: Neural substrates and genetic influences* (pp. 141–160). San Diego, CA: Academic Press.

Mack, N. K. (1995). Confounding whole-number and fraction concepts when building on informal knowledge. *Journal for Research in Mathematics Education, 26,* 422–441. http://dx.doi.org/10.2307/749431

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological methods, 7,* 19–40. http://dx.doi.org/10.1037/1082-989X.7.1.19

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

Mazzocco, M. M. M., & Devlin, K. T. (2008). Parts and 'holes': Gaps in rational number sense among children with vs. without mathematical learning disabilities. *Developmental Science, 11,* 681–691. http://dx.doi.org/10.1111/j.1467-7687.2008.00717.x

McMullen, J., Laakkonen, E., Hannula-Sormunen, M., & Lehtinen, E. (2015). Modeling the developmental trajectories of rational number concept(s). *Learning and Instruction, 37,* 14–20. http://dx.doi.org/10.1016/j.learninstruc.2013.12.004

Meert, G., Grégoire, J., & Noël, M. P. (2009). Rational numbers: Componential versus holistic representation of fractions in a magnitude comparison task. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 62,* 1598–1616. http://dx.doi.org/10.1080/17470210802511162

Meert, G., Grégoire, J., & Noël, M. P. (2010). Comparing 5/7 and 2/9: Adults can do it by accessing the magnitude of the whole fractions. *Acta Psychologica, 135,* 284–292. http://dx.doi.org/10.1016/j.actpsy.2010.07.014

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

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the national mathematics advisory panel*. Washington, DC: U.S. Department of Education.

Ni, Y., & Zhou, Y. (2005). Teaching and learning fraction and rational numbers: The origins and implications of whole number bias. *Educational Psychologist, 40,* 27–52. http://dx.doi.org/10.1207/s15326985ep4001_3

Nylund, K. L. (2007). *Latent transition analysis: Modeling extensions and an application to peer victimization* (Doctoral Dissertation). Retrieved from https://www.statmodel.com/download/nylunddis.pdf

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14,* 535–569. http://dx.doi.org/10.1080/10705510701575396

Nylund-Gibson, K., Grimm, R., Quirk, M., & Furlong, M. (2014). A latent transition mixture model using the three-step specification. *Structural Equation Modeling, 21,* 439–454. http://dx.doi.org/10.1080/10705511.2014.915375

Obersteiner, A., & Tumpek, C. (2016). Measuring fraction comparison strategies with eye-tracking. *ZDM, 48,* 255–266. http://dx.doi.org/10.1007/s11858-015-0742-z

Obersteiner, A., Van Dooren, W., Van Hoof, J., & Verschaffel, L. (2013). The natural number bias and magnitude representation in fraction comparison by expert mathematicians. *Learning and Instruction, 28,* 64–72. http://dx.doi.org/10.1016/j.learninstruc.2013.05.003

Pickering, S., & Gathercole, S. E. (2001). *Working memory test battery for children (WMTB-C)*. London, UK: The Psychological Corporation.

Resnick, I., Jordan, N. C., Hansen, N., Rajan, V., Rodrigues, J., Siegler, R. S., & Fuchs, L. S. (2016). Developmental growth trajectories in understanding of fraction magnitude from fourth through sixth grade. *Developmental Psychology, 52,* 746–757. http://dx.doi.org/10.1037/dev0000102

Schneider, M., & Siegler, R. S. (2010). Representations of the magnitudes of fractions. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1227–1238. http://dx.doi.org/10.1037/a0018170

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6,* 461–464. http://dx.doi.org/10.1214/aos/1176344136

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York, NY: Oxford University Press.

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

Siegler, R. S., & Pyke, A. A. (2013). Developmental and individual differences in understanding of fractions. *Developmental Psychology, 49,* 1994–2004. http://dx.doi.org/10.1037/a0031200

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

Stafylidou, S., & Vosniadou, S. (2004). The development of students' understanding of the numerical value of fractions. *Learning and Instruction, 14,* 503–518. http://dx.doi.org/10.1016/j.learninstruc.2004.06.015

Swanson, J., Schuck, S., Mann, M., Carlson, C., Hartman, K., Sergeant, J., & McCleary, R. (2006). *Categorical and dimensional definitions and evaluations of symptoms of ADHD: The SNAP and SWAN rating scales*. Irvine, CA: University of California.

Thompson, B., & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement, 56,* 197–208. http://dx.doi.org/10.1177/0013164496056002001

Torbeyns, J., Schneider, M., Xin, Z., & Siegler, R. S. (2015). Bridging the gap: Fraction understanding is central to mathematics achievement in students from three different continents. *Learning and Instruction, 37,* 5–13. http://dx.doi.org/10.1016/j.learninstruc.2014.03.002

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *TOWRE: Test of word reading efficiency*. Austin, TX: Pro-Ed.

Van Hoof, J., Verschaffel, L., & Van Dooren, W. (2015). Inappropriately applying natural number properties in rational number tasks: Characterizing the development of the natural number bias through primary and secondary education. *Educational Studies in Mathematics, 90,* 39–56. http://dx.doi.org/10.1007/s10649-015-9613-3

Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis, 18,* 450–469. http://dx.doi.org/10.1093/pan/mpq025

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. San Antonio, TX: The Psychological Corporation.

Wilkinson, G. S., & Robertson, G. J. (2006). *Wide range achievement test (WRAT4)*. Lutz, FL: Psychological Assessment Resources.

Ye, A., & Rinne, L. (July, 2016). Estimating covariance in longitudinal mixture modeling using a three-step approach. Poster presented at the 2016 International Meeting of the Psychometric Society (IMPS), Asheville, NC.

(*Appendices follow*)

Appendix A

Covariate Relationships Between Background Variables and Class Membership at T_1, by Reference Class Parameterization

		Reference class 3 (LNB)			Reference class 2 (SNB)			Reference class 1 (Norm)		
C	Covariate	Est.	*SE*	*p*	Est.	*SE*	*p*	Est.	*SE*	*p*
1	Age	.03	.03	.481	.01	.04	.928			
	Gender	.49	.31	.113	.46	.37	.212			
	Low income	.67	.35	.052	−.10	.41	.815			
	Attentive behavior	−.03	.02	.090	−.05	.02	.002**			
	Verbal ability	−.04	.01	.001**	−.03	.02	.074			
	Nonverbal reasoning	−.16	.08	.039*	.01	.07	.887			
	Reading fluency	.01	.02	.842	−.02	.02	.409			
	Working memory	−.01	.01	.263	.01	.01	.457			
2	Age	.02	.05	.606				−.01	.04	.928
	Gender	.03	.42	.949				−.46	.37	.212
	Low income	.77	.49	.115				.10	.41	.815
	Attentive behavior	.02	.02	.345				.05	.02	.002**
	Verbal ability	−.02	.02	.316				.02	.02	.074
	Nonverbal reasoning	−.17	.10	.064				−.01	.07	.887
	Reading fluency	.02	.03	.414				.02	.02	.409
	Working memory	−.17	.01	.112				−.01	.01	.457
3	Age				−.02	.05	.606	−.03	.04	.481
	Gender				−.03	.42	.949	−.49	.31	.113
	Low income				−.77	.49	.115	−.67	.35	.052
	Attentive behavior				.02	.02	.345	.03	.02	.090
	Verbal ability				.02	.02	.316	.04	.01	.001**
	Nonverbal reasoning				.17	.10	.064	.16	.08	.039*
	Reading fluency				−.02	−.02	.414	−.01	.02	.842
	Working memory				.02	.02	.112	.01	.01	.263

Note. *C* = class number; LNB = larger number bias; SNB = smaller number bias; Norm = Normative.
* $p < .05$. ** $p < .01$.

Appendix B

Correlations Among Predictor Variables in the LTA Model

Predictor	1	2	3	4
1. Attentive behavior				
2. Verbal ability	.361			
3. Nonverbal reasoning	.402	.525		
4. Math achievement	.520	.482	.437	
5. WNLE	−.356	−.383	−.370	−.471

Note. LTA = latent transition analysis; WNLE = whole number line estimation. All correlations are significant at the level $p < .001$.

Received April 22, 2016
Revision received October 28, 2016
Accepted November 14, 2016 ■