Journal of Experimental Psychology: Learning, Memory, and Cognition

Manuscript version of

Do Adults Treat Equivalent Fractions Equally? Adults' Strategies and Errors During Fraction Reasoning

Charles J. Fitzsimmons, Clarissa A. Thompson, Pooja G. Sidney

Funded by:

- Kent State University
- U.S. Department of Education, Institute of Education Sciences

Do Adults Treat Equivalent Fractions Equally? Adults' Strategies and Errors During Fraction Reasoning

Charles J. Fitzsimmons, Clarissa A. Thompson, and Pooja G. Sidney

Author Note

Charles J. Fitzsimmons, Psychological Sciences, Kent State University; Clarissa, A. Thompson, Psychological Science, Kent State University; Pooja G. Sidney, University of Kentucky.

Pre-registrations and data files will be made public upon acceptance for publication and can be viewed here: *https://osf.io/4uygd/*.

Funding: Support for this research was provided in part by the U.S. Department of Education, Institute of Education Sciences grant R305A160295 to Dr. Clarissa A. Thompson and the Kent State University Judie Fall Lasser Award to Charles J. Fitzsimmons.

Conflict of interest: The authors declare that they have no conflict of interest.

Correspondence concerning this article should be addressed to Charles J. Fitzsimmons, Psychological Sciences, Kent State University, Kent, OH 44240.

Email: cfitzsi4@kent.edu

Abstract

Understanding fraction magnitudes is important for achievement and in daily life. However, adults' fraction reasoning sometimes appears to reflect whole number bias and other times reflects accurate reasoning. In the current experiments, we examined how contextual factors and individual differences in executive functioning (Experiment 1), knowledge of fraction equivalence (both experiments), and strategy use (Experiment 2) influenced adults' fraction reasoning. Adults were only biased by fraction components when reasoning about fractions as holistic magnitudes was difficult: when estimating under a time constraint, when estimating fractions with large components, or when comparing fractions close in decimal distance. However, adults' knowledge of fraction equivalence moderated the effects of whole number components on their fraction estimation performance: when modeled at low levels of equivalence knowledge, adults were biased by fraction components when estimating. Adults with more knowledge of fraction equivalence were able to reason about fractions as holistic magnitudes through adaptive strategy choices.

Keywords: whole number bias, fraction reasoning, fraction representations, strategies, adaptive strategy use

Do Adults Treat Equivalent Fractions Equally? Adults' Strategies and Errors During Fraction Reasoning

Imagine you are running late for a friend's surprise party, and you are considering which route to take to get there as fast as possible. Should you take the route that is 15 miles with an average speed limit of 60 miles per hour, or the route that is 10 miles with an average speed limit of 40 miles per hour? If you focus on just the speed, the first route seems best; if you focus on the distance, the second route seems best. As this example illustrates, focusing on any one whole number component would lead to erroneous conclusions because the magnitude of time (e.g., distance divided by speed) is *equivalent* for both routes *(i.e., 15 miles/60 mph = 10 miles/40 mph)*. However, children often only focus on whole number components--numerators and denominators in isolation--when reasoning about rational numbers, thus making *whole number bias* errors (Ni & Zhou, 2005). Even adults appear biased by whole number knowledge in certain contexts, yet appear to accurately reason about fraction magnitudes in others (e.g., Bonato, Fabbri, Umiltà, & Zorzi, 2007; Obersteiner, Van Dooren, Van Hoof, & Vershaffel, 2013; Schneider & Siegler, 2010; Sprute & Temple, 2010).

Understanding fraction magnitudes is difficult for children and adults. Children struggle to order fractions from least to greatest after years of formal fraction instruction (U.S., Department of Education, National Assessment of Educational Progress, 1981, 2008), and community college students struggle with seemingly simple tasks such as selecting the larger of two fractions (e.g., Schneider & Siegler, 2010; Fazio, DeWolf, & Siegler, 2016). This difficulty, likely due to whole number bias, is problematic because understanding the magnitude of fractions is important for success in advanced mathematics (e.g., Booth & Newton, 2012; Fazio, Bailey, Thompson, & Siegler, 2014; Schneider et al., 2018; Siegler et al., 2012), at work

(Handel, 2016), and in everyday life, such as when making medical decisions (e.g., Peters et al., 2006). However, are these difficulties due to a persistent whole number bias, are the errors unique to specific tasks or individuals, and how might various strategies lead to accurate or inaccurate fraction reasoning?

The main goal of the current experiments was to investigate the role of individual differences and contextual factors (i.e., characteristics of the stimuli) in adults' whole number and fraction reasoning. We examined individual differences in knowledge of fraction equivalence and executive functions because both may play important roles in accurate fraction reasoning and help explain the *source* of whole number bias errors. To understand the source of whole number bias errors, we examined how these individual differences *interact* with contextual factors to impact adults' fraction reasoning to evaluate for *whom* whole number components negatively impact fraction reasoning, *when* components bias reasoning, and *how* adults adaptively use strategies to reason about fractions. We examined the who, when, and how of whole number bias errors to evaluate strategic variability accounts of fraction reasoning. In the following sections, we first review the theoretical approaches to whole number bias, then we describe the role of contextual factors, strategies, and individual differences that are important for accurate fraction reasoning.

Theoretical Accounts of Whole Number Bias

Whole number bias is typically characterized by incorrect use of the heuristic *larger components = larger fraction magnitude* (Alibali & Sidney, 2015; Behr, Wachsmuth, Post, & Lesh, 1984; Braithwaite & Siegler, 2017; DeWolf & Vosniadou, 2015; Ni & Zhou, 2005; Stafylidou & Vosniadou, 2004; Vamvakoussi et al., 2012). For example, an adult might reason that traveling 15 miles at a speed of 60 miles per hour is faster than traveling 10 miles at 40 miles per hour because 60 > 40 and 15 > 10, despite both ratios of time being equivalent (15/60 = 10/40 = 0.25). Although the specific source of whole number bias errors is debated, one theoretical account is that children’s initial instruction on whole numbers negatively influences the way they conceptualize subsequently-learned numbers, and these misconceptions--numbers are discrete (have unique successors), have only one symbolic representation, and larger digits represent larger magnitudes--persist into adulthood (Ni & Zhou, 2005). Because fractions differ from whole numbers in a variety of ways, some researchers suggest that children’s initial concept of number must undergo conceptual change to accommodate fraction processing (DeWolf & Vosniadou, 2015; Stafylidou & Vosniadou, 2004; Vamvakoussi, 2015; Vamvakoussi et al., 2012; Vamvakoussi & Vosniadou, 2004; Vamvakoussi & Vosniadou, 2010; Vosniadou, 2014; Vosniadou & Verschaffel, 2004). According to conceptual change theories (e.g., Vosniadou, 2014), children never fully integrate their fraction and whole number knowledge, but instead they continue to process these numbers in distinct ways.

Dual processing accounts. One hypothesis for whole number bias errors in adults is that people process whole numbers intuitively and automatically, but they process fractions effortfully and strategically (Bonato et al., 2010; Kallai & Tzelgov, 2009, 2012; Meert, Grégoire, & Noël, 2009, 2010; Vamvakoussi, 2015; Vamvakoussi & Vosniadou, 2010; Vosniadou, 2014). Thus, the intuitive and effortful processing systems hypothesized to be involved in decision making within dual-processing frameworks are often used to interpret whole number bias errors (Gillard, Van Dooren, Schaeken, & Verschaffel, 2009; Obersteiner et al., 2013; Vamvakoussi, 2015; Vamvakoussi et al., 2012; Van Hoof, Linjen, Verschaffel, & Van Dooren, 2013; Vosniadou, 2014). Within dual-processing accounts, whole number bias might manifest in adults’ longer response times as they attempt to inhibit automatically-activated whole number magnitude knowledge of the fraction components or as increased error rates when they attempt to transform a fraction into a simplified form, such as a decimal or percentage.

The integrated theory of numerical development. Some evidence suggests that fractions and whole numbers *can* be processed more similarly than dual-processing approaches imply. For example, neurological evidence suggests that whole numbers and fractions are processed similarly in the brain (Jacob & Nieder, 2009), and behavioral evidence suggests that people have an ability to process ratios directly without being lured by the magnitude of the whole number components (e.g., Liu, 2017; Matthews & Ellis, 2018; Matthews & Lewis, 2017). Furthermore, according to the integrated theory of fractions and whole number development, one thing that unites both fractions and whole numbers is that their magnitudes can be ordered from smallest to largest on a number line (e.g., Siegler, 2016a; Siegler & Braithwaite, 2016; Siegler, & Lortie-Forgues, 2014; Siegler, Thompson, & Schneider, 2011). Magnitude knowledge improves for increasingly larger numerical ranges and number types (whole numbers, rational numbers, etc.) across developmental time (Opfer & Siegler, 2007; Siegler, 2016a; Siegler, Thompson, & Opfer, 2009; Siegler et al, 2011; Thompson & Opfer, 2010). Although the integrated theory emphasizes the centrality of magnitude knowledge across numerical development, it also emphasizes variability in the types of strategies that people deploy as they encounter fraction problems.

Theories of strategic variability. Variable strategy use is the hallmark of overlapping waves theory (Siegler, 1996, 2005, 2016b). According to overlapping waves theory, strategy selection changes over time with increasing experience, and at any given time-point, individuals use a variety of different strategies to solve problems (Siegler, 1996, 2005, 2016b). Strategic variability accounts of cognition have been extended to explain whole number bias errors. For example, the dynamic strategy choice account of whole number bias (Alibali & Sidney, 2015) is a theoretical framework that acknowledges both the intuitive and strategic processes involved in reasoning about numbers. However, compared to dual-processing approaches, the dynamic strategy choice account places more emphasis on the strategies (both implicit and not) that people use to reason about fractions. Importantly, strategic variability accounts are unique in that they suggest an individual's strategy choice depends on the *interaction* among prior knowledge, experience with similar problems, and characteristics of the problem itself (Alibali & Sidney, 2015; Siegler, 1996, 2005, 2016b). In other words, the strategies and mental processes that people use to reason about fractions will vary *within* individuals on different tasks and problems as well as *between* individuals with different prior knowledge and cognitive abilities. In the next sections, we discuss the contextual factors, such as the size of the whole number numerator and denominator components and time constraints, as well as individual differences, such as equivalence knowledge and executive functioning abilities, that influence adults' fraction reasoning.

Contextual Factors that Influence Whole Number Bias

Component size in comparison tasks. Whole number components influence adults' fraction reasoning in fraction comparison tasks (Bonato et al., 2007; DeWolf & Vosniadou, 2015; Meert et al., 2010; Meert, Grégoire, & Noël, 2012; Obersteiner et al., 2013; Opfer & DeVries, 2008; Vamvakoussi et al., 2012; Zhang, Fang, Gabriel, & Szücs, 2014). In comparison tasks, the larger fraction magnitude in the pair can be comprised of components that are larger and consistent with whole number ordering (e.g., 4/7 > 2/5; 4 > 2 and 7 > 5), or smaller and inconsistent with whole number ordering (e.g., 2/3 > 5/8; even though 2 < 5 and 3 < 8). When comparing fractions with components that are inconsistent with the fraction magnitude, adults tend to have longer response times and higher error rates, which may be due to inhibiting whole number knowledge that is automatically activated by the components (DeWolf & Vosniadou, 2015; Obersteiner et al., 2013; Vamvakoussi et al., 2012). However, the effect of whole number component consistency is only evident when the fraction pairs are close together in decimal distance, which makes comparing holistic magnitudes difficult (DeWolf & Vosniadou, 2015), or when fractions share a common component, and comparing components is adaptive (Obersteiner et al., 2013; Vamvakoussi et al., 2012). For example, comparing only the denominator of unit fractions--fractions with numerators of 1--can lead to fast and accurate responses.

Component size in number line estimation. Even when comparison across components is not possible (e.g., when participants are asked to estimate a single fraction's magnitude on a number line[1]), adults are sometimes biased by the size of fraction components. For example, when adults estimated the location of unit fractions (e.g., 1/60) on a number line ranging from 1/1 on the left to 1/1440 on the right, their estimates corresponded to the size of the denominator, rather than the holistic magnitude of the fractions (Opfer & DeVries, 2008). Unit fractions may be especially likely to elicit whole number reasoning in the estimation task because participants often ignore the common numerator components as they do in fraction comparison tasks (Bonato et al., 2007; Schneider & Siegler, 2010). In the current experiments, we examined whether adults were biased by whole number components when estimating equivalent, non-unit fractions on number lines.

Components and strategic variability. The variable effects of components in fraction comparison and estimation tasks are consistent with strategic variability accounts of cognition.

[1] Although sets of numerator and denominator components cannot be visually compared during completion of the number line estimation task, componential strategies are sometimes reported when participants are asked to describe their strategies after making each estimate (Siegler et al., 2011; Sidney et al., 2018).

For example, in studies of trial-to-trial strategy variability, most adults used six to 11 different strategies across 32 to 48 comparison problems (Fazio et al., 2016; Sidney et al., 2018) and up to seven strategies to estimate fractions on 0-5 number lines (Sidney et al., 2018). Many fraction magnitude comparison strategies involved only comparing components without reference to magnitudes. These componential strategies were adaptive on some trials (when fraction pairs shared a component), yet maladaptive on others (when pairs did not share a component). Other strategies, such as transforming fractions when estimating or referencing magnitudes, demonstrate a holistic representation of fractions as magnitudes. The use of strategies based on holistic magnitudes is adaptive and related to more precise estimation performance for both children and adults (Sidney et al., 2018; Siegler & Thompson, 2014; Siegler et al., 2011). However, the use of magnitude and transformation strategies can be more difficult and time consuming on some trials (e.g., 2/19) than others (e.g., 3/4). Because strategies take different amounts of time to execute, the time allotted will influence strategy selection and performance (Caviola, Carey, Mammarella, & Szücs, 2017; Fazio et al., 2016; Kellogg, Hopko, & Ashcraft, 1999).

Time constraints. Both the dual-processing account and the dynamic strategy choice account of whole number bias suggest that a time constraint should elicit whole number bias errors during fraction reasoning. For example, the dynamic strategy choice account suggests that there are costs and benefits of implementing different strategies; the strategy that leads to the most accurate performance might also be quite time-intensive to execute (Fazio et al., 2016). Therefore, if participants do not have time to execute effective strategies, such as if they are forced to estimate under a time constraint, they will be more likely to make a whole number bias error. Similarly, the dual-processing account points to inhibition of intuitive whole number knowledge and the use of a more analytic processing system--both processes that take additional time--as important ways to combat the whole number bias (e.g., Obersteiner et al., 2013; Van Hoof, Janssen, Verschaffel, & Van Dooren, 2015; Van Hoof, Verschaffel, & Van Dooren, 2015). Thus, to elicit whole number bias, we had adults estimate under a time constraint when they presumably were not given long enough to employ time-intensive, computational strategies or to inhibit automatically-activated knowledge of whole numbers (Campbell & Austin, 2002; Caviola et al., 2017; Fazio et al., 2016; Kellogg et al., 1999). However, some individuals may be able to overcome whole number bias, even if they estimate under a time constraint.

Individual Differences Influencing Whole Number Bias

In the following sections we review how individual differences in knowledge about fraction concepts (e.g., strategies, equivalence) and executive functioning abilities (e.g., inhibition) are important to accurately reason about fraction magnitudes and avoid whole number bias errors.

Equivalence knowledge. Strategic variability accounts predict that prior knowledge may impact adults' ability to accurately reason about fraction magnitudes. Specifically, prior knowledge of fraction equivalence may be important because it requires an understanding that the same magnitude (e.g., 0.5) can be represented by fractions with different whole number components (e.g., 1/2, 15/30, 50%, etc.), yet they are all located at the same place on the number line. Furthermore, adults with more prior knowledge of fraction equivalence may frequently implement strategies that emphasize the holistic magnitude of the fraction, such as transforming one number into the other (e.g., 1/4 = 25%), which may be especially likely to reduce whole number bias errors. Researchers have evaluated how children estimate equivalent fractions one at a time on number lines (Braithwaite & Siegler, 2017), or whether children can match figures (e.g., a circle) to equivalent symbolic fractions (Ni, 2001), but not how equivalence knowledge moderates the effect of whole number components on fraction reasoning. Thus, we directly examined how knowledge of fraction equivalence relates to whole number bias and fraction magnitude understanding. To effectively reason about fraction magnitudes and implement strategies to determine whether two fractions are equivalent also requires executive functioning abilities.

Executive functioning. Executive functioning is thought to be organized into three components: inhibition, updating, and shifting (Miyake et al., 2000) which continue to develop throughout adolescence and serve as a resource to effectively use and mentally manipulate information (Cragg & Gilmore, 2014). Inhibition and updating are proposed to be especially important for understanding fraction magnitudes (Gómez, Jiménez, Bobadilla, Reyes, & Dartnell, 2015; Kolkman, Hoijtink, Kroesbergen, & Leseman, 2013; Siegler & Pyke, 2013). Inhibition is the ability to suppress automatic responses from being executed, and updating is the ability to replace old information in working memory with new, task-relevant information (Miyake et al., 2000). If whole number bias stems from automatically applying whole number knowledge when reasoning about fraction magnitudes, then inhibition of whole number magnitude representations should play an important role in adults' ability to accurately reason about fraction magnitudes (Vamvakoussi et al., 2012; Van Hoof, Janssen, et al., 2015; Van Hoof et al., 2015). If manipulating fractions (e.g., transforming them into equivalent fractions) is important for executing strategies that reduce the likelihood of making whole number bias errors, then updating numerical information in mind should be important for adults' ability to reason about fraction magnitudes. For these reasons, we chose to use domain-specific, numerical measures of inhibition and updating to assess whether inhibition and updating *numerical* information was important for reducing whole number bias errors.

Current Experiments

In the current experiments, we evaluated the role of contextual factors (i.e., fraction stimuli, time constraints) and individual differences (i.e., equivalence knowledge, executive functioning, strategy use) that were hypothesized to impact adults' fraction reasoning when estimating equivalent fractions one at a time. Focusing on equivalent fractions allows us to manipulate component size independent of magnitude to examine whole number bias errors (e.g., 1/2 vs. 15/30 as in Braithwaite & Siegler, 2017). We focus on whole number bias errors during number line estimation to limit the use of componential comparison strategies afforded by the comparison task (Braithwaite & Siegler, 2017). However, we included a fraction comparison task to replicate past research (e.g., DeWolf & Vosniadou, 2015) and evaluate whether a manipulation that may constrain adults' potential strategies during number line estimation subsequently transfers to the fraction comparison task (Siegler & Thompson, 2014).

In line with dual-processing theories and the dynamic strategy choice account of whole number bias, we hypothesized that adults would make whole number bias errors by estimating larger-component fractions as being larger than equivalent, smaller-component fractions, particularly when under a time constraint (H1a). However, the dynamic strategy choice account suggests that whole number components may lead to additional errors beyond the common *larger components = larger fraction magnitude* misconception. For example, strategies such as segmenting the number line may be more difficult and prone to errors when estimating fractions with large (e.g., segmenting the line into 30 parts) compared to small components (e.g., segmenting the line into 2 parts). Furthermore, strategies based on independent whole number components of fractions may lead to less precise estimates because larger whole numbers are represented less precisely along the mental number line (Dehaene, 1992, 2003, 2011; Meert et al., 2009; Meert et al., 2012). Thus, we explored whether component size also affected adults' estimation precision (i.e., PAE).

If imposing a time constraint in the number line task leads to whole number bias errors (i.e., execution of strategies in which the whole number components are processed in isolation), subsequent performance on a comparison task may also be affected. In past work, when children estimated fractions on a number line with landmarks that encouraged componential-based reasoning, they subsequently used componential reasoning when comparing fraction magnitudes (Siegler & Thompson, 2014). If estimating under a time constraint increases componential processing, this may carry over to fraction comparison performance. Thus, we hypothesized that fraction *comparison* accuracy would be lower for adults who estimated under a time constraint as compared to adults who estimated under no time constraint (H1b). We also compared the effects of component consistency and distance on comparison accuracy. We predicted that adults would be less accurate when comparing fractions when the larger fraction magnitude had components that were inconsistent with whole number ordering (H1c), or when the decimal distance between the pair was small (H1d).

We also examined the role of several important individual differences in number line estimation: (1) adults' knowledge of fraction equivalence, (2) inhibition, and (3) updating abilities. The dynamic strategy choice account of whole number bias suggests that individual differences will moderate the effect of whole number components on adults' fraction reasoning in specific contexts (e.g., such as estimating under a time constraint). Thus, we expected individual differences to moderate the interaction between component size and time constraint:

differences in estimates of equivalent fractions with smaller and larger components (i.e., whole number bias) that are made under a time constraint would be smaller for adults with greater understanding of fraction equivalence, better inhibition, or better updating abilities compared to adults with less understanding of equivalence or worse executive functioning abilities (H2).

We replicated and extended our results in a second experiment using a new set of fractions. In Experiment 2, we focused our analyses on individual differences in equivalence knowledge and strategy use. To preview our main findings, we replicated findings from Experiment 1 that demonstrated the importance of equivalence knowledge for reducing the effects of whole number components on estimation precision. Finally, we extended these results to show that estimates were more precise for individuals with more equivalence knowledge, in part, because they used more effective strategies when estimating.

Our combined results across experiments make several novel contributions to the field. First, we directly examined how fraction equivalence knowledge moderates whole number bias errors in adults' number line estimation performance with fractions. To do so, we created a novel two-choice equivalence measure, building on a same/different task (Gabriel, Szücs, & Content, 2013) and a true/false inequality task (Vamvakoussi, Van Dooren, & Verschaffel, 2012). Second, we linked equivalence knowledge to specific strategy use during number line estimation, which provided evidence that theories of strategic variability can explain whole number bias errors during fraction reasoning. Finally, we examined strategy use in the equivalence task, and these reported strategies can be compared and contrasted to strategy use in fraction comparison tasks (e.g., Fazio et al., 2016; Sidney, Thalluri, Buerke, & Thompson, 2018).

Experiment 1

Method

The study was reviewed and approved by the Kent State University Institutional Review Board (protocol #17-432).

Participants

A total of 78 adult participants (*M* = 35.04 *y*, *SD* = 10.25 *y*, range 21 - 69 *y*) met our pre-registered inclusion criteria, and thus had complete data for all tasks[2]. We powered our study to replicate the smallest effect of component size on raw estimates reported in Experiment 2 of Braithwaite and Siegler (2018). We aimed to detect the smallest effect because we thought that adults in the timed condition in our study might perform similarly to the older children in theirs. A power analysis using G*Power (Faul, Erdfelder, Buchner, & Lang, 2009) with 80% power, an alpha of .05, while accounting for 10% attrition indicated we should sample 110 participants; therefore, we sampled 111 total participants from Amazon's Mechanical Turk (MTurk). Participants self-reported being United States citizens currently living in the United States.

We screened participants for inattentive responding using pre-registered inclusion criteria specific to each task (https://osf.io/4uygd/). Participants who did not meet inclusion criteria on number line estimation (*n* = 3), fraction comparison (*n* = 2)[3], fraction equivalence verification (*n* = 2), numerical stroop (*n* = 3), and digit span (*n* = 17), were excluded from analyses involving those tasks. . Five participants were excluded from all analyses for failing to meet inclusion criteria on three out of five tasks (*n* = 2) or for reporting having used a calculator during the experiment (*n* = 3). Despite our attempt to exclude adults who had recently taken similar surveys

[2] We included participants who completed all tasks in order to make results comparable across all analyses, however this differed slightly from our pre-registered inclusion criteria. The conclusions and major findings are the same using either inclusion criteria.

[3] One participant was excluded from the comparison task for only answering one problem correctly.

from our research team, 13 adults with such prior experience participated and were excluded. For detailed information about inclusion criteria, please see our pre-registration link.

Of the 78 participants in the final data set, 30 were female and 47 were male; one participant did not report gender. Most participants were white (*n* = 55; 70.5%), had a college degree (*n* = 34; 43.6%) or completed some college (*n* = 20; 25.6%), and were employed for wages (*n* = 51; 65.4%). See Table 1 for all demographic information.

Table 1.

Demographics from Experiments 1 and 2.

Demographics		
	Frequency (percent)	
Race	E1	E2
White	55 (70.5)	64 (64)
Black or African American	8 (10.3)	10 (10)
American Indian or Alaskan Native	4 (5.1)	2 (2)
Asian	4 (5.1)	13 (13)
Hispanic or Latino	1 (1.3)	5 (5)
Multiple races	5 (6.4)	4 (4)
Did not report	1 (1.3)	2 (2)
Education		
Completed high school	14 (17.9)	12 (12)
Post high school other than college	1 (1.3)	3 (3)
Some college	20 (25.6)	24 (24)
College graduate	34 (43.6)	47 (47)
Postgraduate	8 (10.3)	11 (11)
I'd rather not report	1 (1.3)	3 (3)

Design and Tasks

All participants estimated equivalent fraction magnitudes that increased in component size (smallest, small, large, and largest); approximately half of participants (*n* = 49) were randomly assigned to complete the number line estimation task under a three second per trial time constraint (henceforth timed condition), and the other half (*n* = 44) completed estimates at their own pace (henceforth untimed condition). Thus, we tested a 2 (between: timed vs. untimed) x 4 (within: smallest, small, large, largest component) design.

Number line estimation. Participants indicated the location of a fraction on a number line with 0 on the left and 1 on the right by either clicking on the line or sliding an icon to a particular location on the line. Forty-four fractions with 11 magnitude values ranging from 0.2 to 0.83 (1/5, 2/9, 1/4, 1/3, 3/7, 1/2, 5/9, 2/3, 3/4, 4/5, and 5/6) were adopted from Experiment 2 of Braithwaite and Siegler (2017). Each fraction was presented with four different numerator and denominator component sizes (see Appendix A). For example, fractions with magnitudes equivalent to 0.5 were 1/2, 2/4, 12/24, and 15/30. The smallest component fractions were always in their simplified form (e.g., 1/2). Braithwaite and Siegler (2017) multiplied the simplified fractions by a number, such as 2/2 or 3/3, to create equivalent, non-simplified small component fractions. They multiplied each component in the simplified fractions by a larger number (e.g., 12/12 or 15/15) to create the large (e.g., 12/24) and largest (e.g., 15/30) component equivalent fractions. Thus, component size is operationalized as relative to the smallest-component, equivalent fraction. We changed two non-simplified small component fractions from the original stimuli to further clarify the distinction between small and large component fractions.

As in Braithwaite and Siegler (2017), we compared participants' raw estimates of small and large component fractions to evaluate the whole number bias. In line with prior research

(Siegler & Booth, 2004; Siegler et al., 2011), we also assessed magnitude knowledge by calculating percent absolute error (PAE) for each trial and averaging across all trials for each participant. Percent absolute error was calculated by taking the absolute value of the decimal distance between the actual fraction magnitude and the participant's estimate then dividing the result by the numerical range. PAE is expressed as a percentage by multiplying the result by 100 and is averaged across trials for each participant:

$$PAE = ((|\text{actual - estimate}|) / \text{numerical range}) * 100$$

Participants completed number line estimates in one of two quasi-randomized orders so that participants did not estimate two equivalent fractions in a row (e.g., 15/30 was never presented after 1/2).

Confidence judgments. After participants estimated each fraction on the number line, they immediately rated their confidence on a 4-point scale (adopted from Wall, Thompson, & Morris, 2015). Analyses pertaining to confidence judgments were not central to our current hypotheses so they will not be reported.

Fraction comparison. Participants selected which of two fractions was larger in magnitude across 20 trials (see Appendix A). All fraction magnitudes were less than one, with both fraction magnitudes in some pairs less than (e.g., 12/28 vs. 2/5) or greater than (e.g., 11/14 vs. 15/21) 0.5, or with one of the magnitudes less than and the other greater than 0.5 (e.g., 13/29 vs. 10/12). The fraction with the larger magnitude was presented equally on the right and left side of the screen, and all items were presented randomly to each participant. The instructions emphasized speed and accuracy and discouraged calculator use.

We created fraction comparison pairs that varied in the consistency of the whole number components and the numerical distance between fractions to replicate prior work on fraction comparison (DeWolf & Vosniadou, 2015). For *consistent* pairs, participants comparing holistic magnitudes would make the same choice as participants relying on the "larger components, larger fraction magnitude" heuristic: the fraction with the larger magnitude had larger numerator and denominator components than the fraction with the smaller magnitude (e.g., 14/18 > 2/5). For *inconsistent* pairs, using this componential strategy would lead to incorrect responses: the fraction with the larger magnitude had *smaller* numerator and denominator components than the fraction with the smaller magnitude (e.g., 6/7 > 12/29). No fractions within pairs had equivalent numerators or denominators, as this has been shown to elicit componential comparisons (Bonato et al., 2007; Schneider & Siegler 2010). We created pairs with either a small (i.e., < .14) or large (i.e., > .37) distance between them. We examined accuracy (out of five trials) for inconsistent small distance trials, inconsistent large distance trials, consistent small distance trials, and consistent large distance trials. Because adults use a variety of strategies to compare fractions, and these strategies take different amounts of time to execute, we chose not to evaluate response times (for a discussion of this issue, see Fazio et al., 2016).

Fraction equivalence knowledge. Participants verified whether 66 fraction equivalence statements (e.g., 1/2 = 15/30) were true or false (see Appendix A). Fraction stimuli were drawn from the number line estimation task. Each simplified fraction was paired with an equivalent small component fraction (e.g., 1/2 = 2/4) and with the largest component fraction (e.g., 1/2 = 15/30); the largest component fractions were paired with other small component fractions (e.g., 2/4 = 15/30). We created false statements using the same component size pairings (e.g., 6/8 = 1/2, 1/2 = 12/36, and 2/4 = 12/16), while controlling for the decimal distance between each fraction because distance impacts fraction comparison speed and accuracy (DeWolf &

Vosniadou, 2015). Approximately half of the fractions in our false statements differed by a small distance (<.14) and the other half differed by a large distance (>.37).

As with fraction comparison, the fraction with the larger magnitude was presented equally on the right and left side of the screen (for false items), the fraction with larger components was presented equally on the right and left side of the screen, and all items were presented randomly to each participant. Equivalence knowledge was measured as the percentage of correct responses out of 66 possible trials.

Inhibition. Participants’ inhibition was measured with a numerical stroop task in which they viewed a fixation cross for 300 *ms*, then had 2,000 *ms* to select which of two digits was physically larger while ignoring numerical magnitude (i.e., inhibit whole number magnitude knowledge). We adopted digit pairs with a magnitude distance of one or five (distance 1: 1 vs. 2, 3 vs. 4, 6 vs. 7, 8 vs. 9; distance 5: 1 vs. 6, 2 vs. 7, 3 vs. 8, 4 vs. 9) using eight single digits from Dadon and Henik (2017). Digit pairs were presented with a physical ratio (smaller font size [56 point Arial] divided by larger font size [67 point Arial]) of 0.84, which produced the longest reaction times in a prior study (Leibovich, Diesendruck, Rubinsten, & Henik, 2013). Participants completed a random presentation of 112 digit pairs consisting of 32 congruent (numerical and physical size in the same direction, e.g., 1 2), 32 incongruent (numerical and physical size in opposite directions, e.g., 1 2), and 48 neutral (only manipulate the task-relevant dimension of physical size, e.g., 1 1) trials. The larger digit appeared equally on the right and left side of the screen.

We measured participants’ numerical inhibition ability with their average response times (RT) to correct responses on incongruent trials (as in Gómez et al., 2015; Kaufmann et al., 2005; Leibovich et al., 2013). We chose to use only the average response time to incongruent trials because of inherent problems with the reliability of difference scores as individual difference measures (Cronbach & Furby, 1970; Edwards, 1995, 2001), and because the use of RT from incongruent trials only has been used in prior work (e.g., Gómez et al., 2015). However, we confirmed that adults' response times were longer on incongruent (*M* = 755ms, *SD* = 181ms) than on neutral (*M* = 640ms, *SD* = 166ms), $t(77) = 13.92$, $p < .001$, and congruent trials (*M* = 614ms, *SD* = 173ms), $t(77) = 14.82$, $p < .001$. We also confirmed that adults' response times were longer on neutral than congruent trials, $t(77) = 3.72$, $p < .001$. These results indicate adults are inhibiting the automatically activated knowledge of the whole number magnitude during incongruent trials to respond correctly.

Updating. Participants' updating was measured with a digit span task in which they viewed a fixation cross for 2,000 *ms*, then viewed a series of digits, one at a time for 2,000 *ms* each. Participants were instructed to recall, in order, the last four digits presented when a "????" appeared on the screen (see Siegler & Pyke, 2013 for a similar letter span task). The instructions encouraged participants to practice saying *only* the last four digits out loud as each new digit appeared. Each sequence of five, seven, nine, or 11 digits was randomly presented three times for a total of 12 trials. Digits within test sequences were presented in a fixed order, and answers always contained a unique series of digits (i.e., with no repeating numbers: 7942). We measured participants' updating ability as the percentage of correctly recalled digits out of 48.

Fraction familiarity. Participants indicated how familiar they were with 44 fractions on a six-point scale from "not familiar at all" to "very familiar" (see Dumas, Johnson, & Lynch, 2002 for a similar word-familiarity scale). Analyses pertaining to familiarity judgments were not central to our current hypotheses so they will not be reported.

Demographic information. Participants reported their age, gender, race/ethnicity, level of education, their scores on the math section of the SAT/ACT, the country in which they completed their primary education, the most advanced mathematics course they completed, their employment status, and the device they used to complete the survey.

Procedure

Participants accepted the Human Intelligence Task (HIT) through Amazon's Mechanical Turk and then were directed to a link to complete our tasks on the Qualtrics website. On average, participants took 46 minutes to complete the survey (overall range = 24.65 - 129.97 minutes; timed condition: 24.65 – 98.97 minutes; untimed condition: 25.32 – 129.97 minutes). First, participants completed the number line estimation task, then the fraction comparison task, and then the fraction equivalence task. Participants then completed the numerical stroop and digit span tasks in a counterbalanced order. Finally, all participants rated the familiarity of fractions and answered demographic questions. Participants were compensated $6.00.

Participants were randomly assigned to complete number line estimates either under a three second per trial time constraint or at their own pace. This time limit was determined using half of the mean response time it took pilot participants to solve the number line estimation problem under no time constraint, as has been done in past work (Crescenzi, Kelly, & Azzopardi, 2015). Furthermore, the three second time limit that we implemented was shorter than the average 5.95 seconds it took adults from a highly selective university to compare fractions. Three seconds was also notably shorter than the average response time for high performing (9.33s) and low-performing community college students (10.91s) reported in Fazio et al. (2016). If participants exceeded the three second time constraint, the screen flashed red and informed participants to, "Please respond within the time limit," but the trial only advanced after participants made their estimate. Thus, there was no missing data for any participant in the number line task. Participants exceeded the three second time limit on 16.5% of trials. Percent absolute error (PAE) for estimates made over the time limit (*M* = 10.27%, *SD* = 9.09%) did not differ from estimates made within the time limit (*M* = 8.93%, *SD* = 8.11%), $t(38) = 1.38$, p = .177.

Results

Preliminary Analyses

First, we confirmed that random assignment to number line estimation condition resulted in equivalent groups by examining demographics and other individual differences. Participants in the timed and untimed conditions did not differ in age, $t(75) = 0.53$, $p = .598$ (one participant did not report age), gender, $X^2(2, N = 78) = 2.02$, $p = .364$, performance on the equivalence task, $t(76) = .47$, $p = .639$, inhibition task, $t(76) = 1.21$, $p = .231$, or updating task, $t(76) = 0.17$, p = .865. Thus, it appeared that random assignment was effective. We confirmed that there were no differences in overall percent absolute error (PAE) on number line estimates between the two randomized orders of fractions in the timed, $t(39) = 0.24$, $p = .809$, and untimed, $t(35) = 0.75$, p = .457, conditions. We collapsed the two orders of fractions within each condition for all analyses.

As can be seen in Table 2, we replicated prior work showing that number line estimation and fraction comparison performance were related (e.g., Fazio et al., 2014; Laski & Siegler, 2007; Siegler et al., 2011). Furthermore, our measure of fraction equivalence knowledge related to estimation and comparison performance. However, inhibition only related to performance in the comparison task, while updating related to performance in the number line estimation, comparison, and equivalence tasks.

Table 2.

Descriptive statistics and correlations between tasks.

	N	α	Mean (SD)	1	2	3	4	5
1. *Number Line Magnitude PAE*[a]	78	.964	8.4% (7.7%)	-	-	-	-	-
2. *Comparison Accuracy*	78	.728	76.2% (16.7%)	**-.568****	-	-	-	-
3. *Equivalence Accuracy*	78	.802	84.1% (12.0%)	**-.536****	**.497****	-	-	-
4. *Numerical Stroop RT (inhibition)*	78	.950[b]	755ms (181ms)	.178	**-.277***	-.121	-	-
5. *Digit Span Accuracy (updating)*	78	.847	83.3% (17.7%)	**-.370****	**.361****	**.327****	.004	-

Notes. * $p < .05$, ** $p < .01$ [a]PAE is a measure of error in which higher percentages indicate less precise estimates. [b]Cronbach's alpha for numerical stroop was calculated based on response times to correct responses on incongruent items only using the *Psych* package from R with pairwise deletion (Revelle, 2018).

Analysis Plan

We predicted that adults would estimate fractions with larger components as larger than equivalent fractions with smaller components, especially when estimating under a time constraint (H1a). However, we hypothesized that having more knowledge of fraction equivalence or better executive functioning abilities would reduce the effects of a time constraint on estimates of equivalent fraction magnitudes (H2). Across all number line analyses, we tested our hypotheses in between-within, general linear models (GLMs) with condition (untimed vs. timed) as a between-subjects factor, component size (smallest, small, large, and largest) as a within-subjects factor, and equivalence knowledge, updating, and inhibition abilities as concomitant variables[4].

[4] This analysis differs from our pre-registered difference score approach, but the conclusions are similar when we do not collapse categories, but instead analyze component size as a four-level independent factor. We chose to analyze component size as a four-level independent factor to

Note, we label these as concomitant variables rather than covariates because we are testing the interaction between the within-subjects factor, the between-subjects factor, and continuous predictors (see Judd, Kenny, & McClelland, 2001; Maxwell & Delaney, 2004). We explored significant interactions with individual difference measures at high (+1 SD from mean) and low (-1 SD from mean) levels of the continuous individual difference variable in the timed and untimed conditions using planned polynomial contrasts on estimated marginal means. These follow-up planned contrasts are equivalent to simple slope tests. Simple slope tests for marginal interactions are reported in the Supplemental Analyses File. Bonferroni corrections were used in post-hoc, pairwise comparisons to adjust for multiple tests. When sphericity was violated, we report Greenhouse-Geisser corrected degrees of freedom.

Individual Differences in Number Line Estimation

Raw estimates. We conducted a 4 (component size: smallest, small, large, largest) x 2 (condition: untimed vs. timed) between-within mixed GLM with equivalence, inhibition, and updating as concomitant variables on raw number line estimates. The results of the full model can be seen in Table B1. Tests of within-subjects effects revealed no main effect of component size on raw estimates, $F(2.085, 145.918) = 0.52$, $p = .604$, but component size interacted with condition, $F(2.085, 145.918) = 7.71$, $p = .001$, $\eta_p^2 = .10$. There were no other within-subjects, two-way interactions. In line with our second hypothesis (H2), the interaction between component size and condition was qualified by a three-way interaction with equivalence knowledge, $F(2.085, 145.918) = 8.04$, $p < .001$, $\eta_p^2 = .10$, and a marginal three-way interaction with inhibition, $F(2.085, 145.918) = 2.55$, $p = .079$, $\eta_p^2 = .04$.

---

evaluate whether whole number bias increases as the size of components increases and to reduce the number of analyses necessary to test our hypotheses.

Component size by condition. We conducted follow-up pairwise comparisons examining the effects of component size in the untimed and timed condition on the estimated marginal means of raw estimates. Adults estimated the largest-component fractions as 0.023 larger than large-component fractions when they estimated under a time constraint (p = .008). However, no other comparisons in the timed or untimed conditions were significant (all ps > .14). Thus, in contrast to our first hypothesis (H1a), participants in our study rarely demonstrated whole number bias errors in their raw estimates of fractions, even when under a time constraint.

The moderating role of equivalence knowledge. To explore the three-way interaction between component size, condition, and equivalence knowledge, we tested the effect of component size in each condition when equivalence knowledge was higher (+ 1 SD) or lower (-1 SD). As can be seen in Figure 1, all pairwise comparisons of estimated marginal means were non-significant in the untimed condition.

In the timed condition, multivariate tests revealed an effect of component size when equivalence knowledge was higher, $F(3, 68) = 4.50$, $p = .006$, $\eta_p^2 = .17$, and lower, $F(3, 68) =$ 4.43, $p = .007$, $\eta_p^2 = .16$. However, the pattern of pairwise comparisons on estimated marginal means is less clear. As can be seen in Figure 1, when modeled at lower equivalence knowledge, adults estimated the smallest component fractions as *larger* than fractions with small and large components. In contrast, when modeled at higher levels of equivalence knowledge, adults were only marginally affected by component size and estimated the smallest component fractions as *smaller* than other equivalent fractions.

Thus, whether whole number components influenced adults' fraction reasoning depended on the difficulty of the task and their prior knowledge of fraction equivalence. Only when equivalence knowledge was lower were adults influenced by the size of fraction components when estimating under a time constraint. Furthermore, the effect of components was in contrast to how whole number bias is typically characterized. Adults with less equivalence knowledge may have made errors when attempting to mentally simplify a larger component fraction, or they may have used the heuristic that *smaller components = larger magnitudes* without taking into consideration the size of the numerator. These data are consistent with strategic variability accounts of fraction reasoning (Alibali & Sidney, 2015): when under restricted time, adults' performance suggested that which strategies and heuristics they used depended on their knowledge of fraction equivalence.

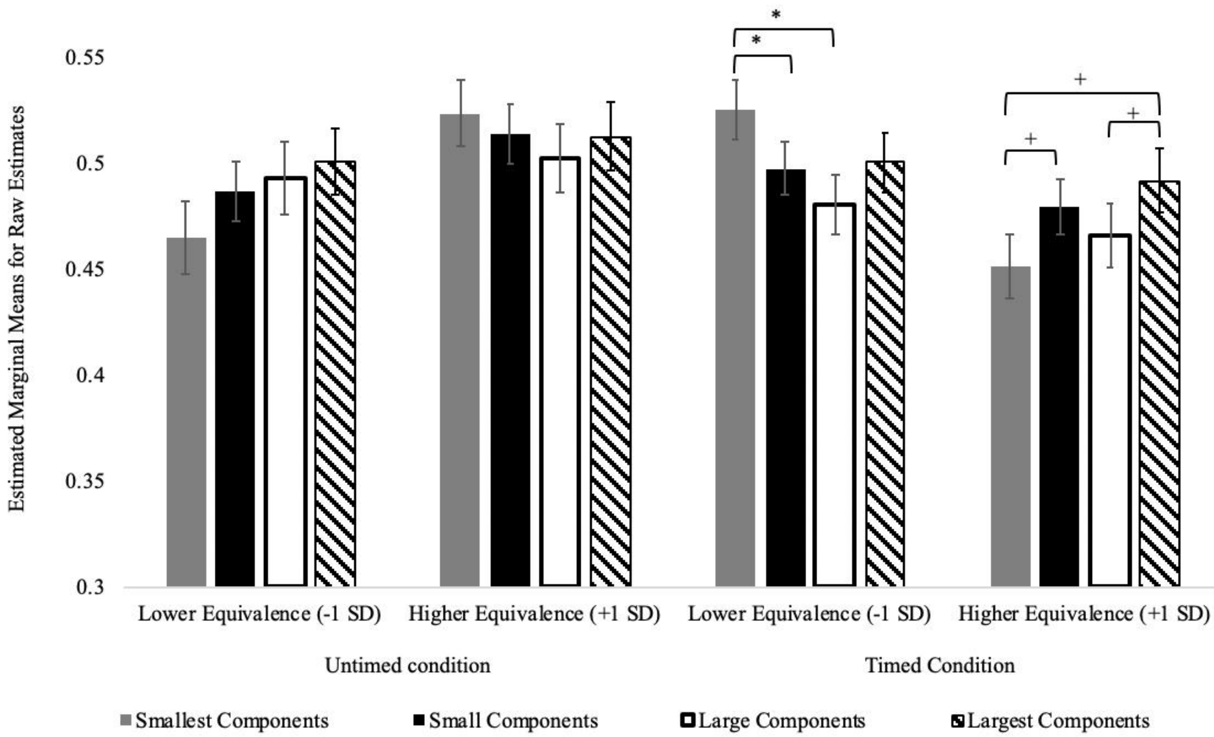

Figure 1. Three-way interaction between condition, component size, and equivalence knowledge on average raw estimates. *Note*. All estimates are estimated marginal means at high (+ 1 SD) and low (- 1 SD) levels of equivalence knowledge while controlling for mean values of inhibition and updating. $^{*}p < .05$, $^{+}p < .1$.

Estimation precision. We ran a parallel mixed 4 (component size: smallest, small, large, largest) x 2 (condition: untimed vs. timed) between-within GLM with equivalence, inhibition, and updating as concomitant variables on estimation precision (PAE). Within-subjects effects did not reveal a significant effect of component size on PAE, $F(2.742, 191.93) = 1.49$, $p = .223$. However, component size interacted with condition, $F(2.742, 191.93) = 3.58$, $p = .018$, $\eta_p^2 = .05$, and this interaction was qualified by a three-way interaction with equivalence knowledge, $F(2.742, 191.93) = 4.28$, $p = .008$, $\eta_p^2 = .06$. There was a marginal three-way interaction between component size, condition, and updating, $F(2.742, 191.93) = 2.26$, $p = .088$, $\eta_p^2 = .03$, but no other interactions were significant (see Table B2 for the full model).

The moderating role of equivalence knowledge. When modeled at higher (+ 1 SD) levels of equivalence knowledge, there was no multivariate effect of component size on estimation precision for adults in the untimed, $F(3, 68) = 0.73$, $p = .538$, or timed condition, $F(3, 68) = 1.13$, $p = .344$.

When modeled at lower equivalence knowledge (-1 SD), there was a significant multivariate effect of component size in both the untimed, $F(3, 66) = 7.49$, $p < .001$, $\eta_p^2 = .25$, and timed conditions, $F(3, 68) = 6.22$, $p = .001$, $\eta_p^2 = .22$. As can be seen in Figure 2, estimated marginal means were higher (less precise) for fractions with the largest components compared to fractions with smaller components in both conditions. Thus, regardless of timing condition, the effect of component size on estimation precision depended on adults' prior knowledge of fraction equivalence. When equivalence knowledge was higher, adults estimated fractions based on their holistic magnitude, regardless of the size of components or time constraints.

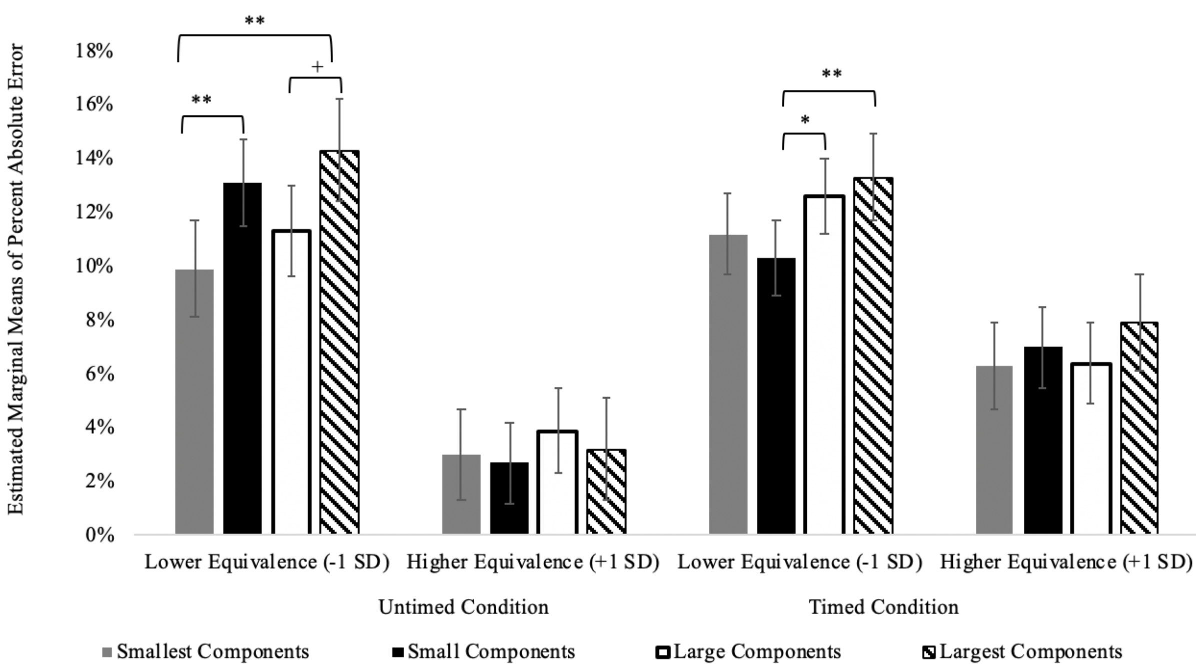

Figure 2. Three-way interaction between component size, condition, and equivalence knowledge on percent absolute error (PAE) while controlling for updating and inhibition. Effects are based

on estimated marginal means at lower (-1 SD) and higher (+ 1 SD) levels of equivalence knowledge. *Note*. Higher PAE relates to less precise estimates. * $p < .05$. ** $p < .01$ $^{+}p < .1$.

Fraction Comparison Performance

Accuracy. We hypothesized that adults would be less accurate when comparing fractions after making number line estimates under a time constraint (H1b), when comparing fractions with components inconsistent with whole number ordering (H1c), and when the distance between the pair was small (H1d). To evaluate our hypotheses, we conducted a mixed ANOVA with condition (untimed vs. timed) as a between-subjects factor, and consistency (consistent vs. inconsistent) and distance (small vs. large) as within-subjects factors.

Components only influenced adults’ fraction reasoning when it was difficult to compare fractions based on holistic magnitudes (see Table B3). Within-subjects effects revealed a significant effect of consistency, $F(1, 76) = 6.86$, $p = .011$, $\eta_p^2 = .08$, distance, $F(1, 76) = 43.42$, $p < .001$, $\eta_p^2 = .36$ and an interaction between consistency and distance, $F(1, 76) = 21.95$, $p < .001$, $\eta_p^2 = .22$. Adults were *less* accurate at comparing *consistent* ($M = 71.7\%$, $SE = 2.4\%$) than inconsistent fraction pairs ($M = 81.0\%$, $SE = 2.7\%$), and when comparing small- ($M = 68.2\%$, $SE = 1.8\%$) than large-distance ($M = 84.7\%$, $SE = 2.7\%$) pairs. The interaction revealed that the effect of whole number component consistency depended on the distance between the fractions.

The effect of consistency was only significant when adults compared small-distance fraction pairs, $F(1, 76) = 14.25$, $p < .001$, $\eta_p^2 = .16$. This surprising result, which replicates prior work with Greek participants (DeWolf & Vosniadou, 2015), might be due to adults comparing fractions by selecting the fraction with the smaller within-fraction numerator and denominator difference. This componential heuristic has been reported in past work examining fraction comparison strategy use (Fazio et al., 2016; Sidney et al., 2018). In our study, this heuristic would lead to correct responses on 100% of trials when comparing inconsistent fraction pairs. However, when comparing consistent fraction pairs that were close in decimal distance, this heuristic would lead to the incorrect response on four out of five trials.

Between-subjects effects did not reveal an effect of condition, $F(1, 76) = 1.18$, $p = .280$, $\eta_p^2 = .02$. Adults in the untimed condition had an average overall accuracy (M = 78.4%, SE = 2.7%) similar to adults in the timed condition (M = 74.3%, SE = 2.6%). Condition did not moderate the effect of consistency, distance, or the interaction between consistency and distance (all ps > .3).

Experiment 1 Discussion

In line with strategy variability accounts of whole number bias (e.g., Alibali & Sidney, 2015), adults' performance when estimating and comparing fractions suggested they adapted their strategy choices to the task. When it was difficult to reason about fractions as holistic magnitudes (e.g., when under a time constraint or when comparing fractions close in distance), adults fell back on componential-based heuristics. By contrast, adults' fraction reasoning under a time constraint did not reflect whole-number-based strategies. One possibility is that the time constraint was not constraining enough to lead to whole number bias errors. However, this seems unlikely. The average number line estimation trial-level response time was significantly longer in the untimed (M = 14.97s, SD = 17.74s) than timed (M = 3.26s, SD = .35s) conditions, $t(36.026) = 4.01$, $p < .001$. This was also the case for estimates of fractions from each component size bin (all ps < .01). Furthermore, as mentioned earlier, the time constraint we imposed in the current study is much shorter than the time it takes college students enrolled in a highly selective university to compare fraction magnitudes (Fazio et al., 2016).

In line with our second hypothesis (H2), only adults with less knowledge of equivalence[5] were influenced by components when reasoning about fractions one at a time in the number line task. Adults with less knowledge of equivalence estimated larger-component fractions less precisely than smaller-component fractions in both conditions and estimated fractions with *smaller components* as larger in magnitude in the timed condition. Adults with more knowledge of equivalence estimated smaller- and larger-component fractions with similar precision in both conditions, estimated fractions similarly in the untimed condition, and were only marginally affected by components in the timed condition.

The moderating role of equivalence knowledge during fraction estimation suggests that individual differences in knowledge impact 1) what strategies are selected and 2) how effectively they are implemented. For example, adults with more equivalence knowledge may have adaptively used strategies in both conditions but did not have enough time to effectively implement them in the timed condition. In contrast, when it was difficult to reason based on holistic magnitudes, adults with less equivalence knowledge were likely falling back on less adaptive heuristics, such as *smaller components = larger magnitude*. This heuristic aligns with a part-whole interpretation of fractions: when the denominator is smaller, there are fewer, but larger, parts (i.e., a fraction with a denominator of 5 has larger parts than a fraction with a denominator of 19). However, if the numerator is also small (e.g., 2), the fraction represents fewer *parts* of the *whole* (i.e., 2/5 is only 2 out of a possible 5 parts). This part-whole interpretation is adaptive when reasoning about unit fractions, but not all fractions. Thus, knowledge of fraction equivalence was important for adaptively selecting strategies when reasoning about fraction magnitudes, in line with overlapping waves theory (Siegler, 1996, 2005,

[5] Note that we use a person-centered interpretation of our results. However, these findings are based on modeled effects and we did not group participants based on their performance on the equivalence task.

2016b) and the dynamic strategy choice account of whole number bias (Alibali & Sidney, 2015). However, we did not directly examine strategy use in Experiment 1, which was the primary motivation for conducting Experiment 2.

Experiment 2

In Experiment 2, we examined *why* individual differences in equivalence knowledge moderate the effects of contextual factors on adults' fraction reasoning by examining self-reported strategy use. In addition to examining strategy use, we investigated the role of task order on number line estimation performance. We explored whether recent experiences can *facilitate* accurate fraction reasoning.

Recently activated knowledge can help or hinder performance and strategy use (Crooks & Alibali, 2013; McNeil & Alibali, 2005; Sidney & Alibali, 2017). For example, adults used less effective strategies to solve mathematical equivalence problems (e.g., 3 + 4 + 6 = 3 + __) when irrelevant knowledge was activated (Crooks & Alibali, 2013; DeCaro, 2016; McNeil & Alibali, 2005). Research involving children suggests prior experiences can also help performance. For example, children who had their knowledge of *whole number* division activated first provided more conceptually accurate representations of fraction division on a following fraction division task compared to children whose knowledge of fraction addition was activated (Sidney & Alibali, 2017). Thus, we tested whether activating adults' knowledge of equivalence, by having participants complete the equivalence task before number line estimation, would reduce whole number bias effects.

We hypothesized that component size and equivalence knowledge would interact: adults with more knowledge of equivalence would estimate small- and large-component fractions with similar precision, but adults with less equivalence knowledge would be less precise when estimating large- relative to small-component fractions (H1). Given the lack of a significant effect of component size on raw estimates in the untimed condition from Experiment 1, we did not expect significant differences in raw estimates of fractions with smaller or larger components. However, the differences in component size between small- and large-component fractions using stimuli from Braithwaite and Siegler's (2017) study were always relative to the smallest-component fractions. Indeed, some small-component fractions (e.g., 10/12) had similar components as large-component fractions of different magnitudes (e.g., 9/12). Thus, in Experiment 2 we distinguished component size by making all large component fractions have double digit components.

Of primary interest, we examined the strategies that adults used and related the frequency of their strategy usage to estimation precision. Self-reported strategies have been successfully used in prior research to gain insight into the strategies children and adults use to reason about other numerical tasks such as multiplication, addition, and fraction comparison or estimation (e.g., Campbell & Alberts, 2009; Fazio et al., 2016; LeFevre, Bisanz, et al., 1996; LeFevre, Sadesky, & Bisanz, 1996; Sidney et al., 2018). In the current experiment, solving problems in the number line estimation and fraction equivalence tasks takes long enough to form a representation in short-term (or working) memory, and the representations are active long enough that the nature of how these representations have been manipulated (i.e., the strategies) can be reported. Thus, these tasks meet Ericsson and Simon's (1980, 1984, 1993, 1998) requirements for providing valid strategy reports. However, we acknowledge that there are some limitations to self-reporting strategies (e.g., Smagorinsky, 1998; Thevenot, Castel, Fanget, & Fayol, 2010; Wilson & Nisbett, 1978). For example, individuals' behavioral performance sometimes suggests they are using strategies that they do not self-report when prompted (Thevenot et al., 2010) or reporting strategies may change performance, strategies, or thinking (e.g., Smagorinsky, 1998; c.f., Ericsson & Simon, 1998 or Smith & Miller, 1978 for counter arguments). Although self-reporting strategies may have limitations, this methodology is one way to evaluate predictions made by the dynamic strategy choice account of whole number bias. By having adults provide strategy reports, we do not have to rely solely on experimenter-generated inferences about why people performed the way they did; they tell us explicitly. The combination of self-reported strategies with the behavioral data from Experiment 1 provides converging evidence of the processes associated with adults' fraction reasoning.

In Experiment 2, we predicted that use of adaptive magnitude strategies would relate to more precise estimates of fractions on number lines (H2a) and higher overall accuracy on the equivalence task (H2b). We also explored whether knowledge of fraction equivalence related to more precise number line estimates through the use of adaptive strategies.

Method

The study was reviewed and approved by the Kent State University Institutional Review Board (protocol #17-432).

Participants

A total of 100 participants (M_{age} = 34.94 *y*, *SD* = 11.40 *y*, range 19 - 71 *y*; 42% female) met pre-registered inclusion criteria and were included in the final data set (see Table 1). We recruited 112 adult participants who self-reported being at least 18 years old, United States citizens, and currently living in the United States. As in Experiment 1, we powered to detect a small effect with 80% power and the potential for 10% attrition (Faul et al., 2009). We powered to detect a small effect to rule out the possibility that we did not detect an effect of component size in the untimed condition in Experiment 1 due to power. The same pre-registered inclusion criteria from Experiment 1 was implemented in Experiment 2. There were 12 participants who were excluded from analyses: seven from number line estimation, one from fraction equivalence, and seven for reporting having used a calculator (there is some overlap between calculator use and other exclusions).

Design and Tasks

The design of Experiment 2 was similar to that of Experiment 1. All participants estimated equivalent fraction magnitudes with small or large components under no time constraint. Approximately half of the participants were randomly assigned to complete the number line estimation and equivalence tasks in either the activated (n = 45) or non-activated (n = 55) condition. Participants in the activated condition completed the fraction equivalence task before number line estimation, and those in the non-activated condition completed tasks in the opposite order. Thus, we tested a 2 (between: activated vs. non-activated) x 2 (within: small vs. large component) between-within design with equivalence knowledge as a concomitant variable.

Number line estimation. This task was similar to Experiment 1, but participants estimated the magnitude of a different set of 12 fractions with six magnitude values ranging from 0.167 to 0.857 (see Appendix A). We adopted a subset of simplified, small component fractions from experiment one of Braithwaite and Siegler (2017). We created equivalent large component fractions with double-digit numerators and denominators because the largest differences in estimation precision in Experiment 1 were between the largest and smallest component fractions. As in Experiment 1, fractions were presented in one of two pre-randomized orders so that no equivalent magnitude was presented back-to-back. We analyzed both raw estimates and percent absolute error to evaluate the effects of component size on adults' fraction reasoning.

Fraction equivalence verification. This task was the same as in Experiment 1. Participants completed nine true and nine false equivalence statements using fraction stimuli drawn from the number line task. Because all small component fractions in the number line task were in their simplified form, the six original true statements could be correctly solved using a componential equivalence strategy: do the components of one fraction evenly go into to the components of the paired fraction (e.g., the numerator, 1, goes into the numerator, 12, evenly and the denominator, 6, goes into the denominator, 72, evenly when evaluating 1/6 = 12/72). Thus, we created three additional true statements using non-simplified fractions equivalent to three of the original magnitudes (1/6, 4/9, 6/7) that were not included in the number line task. These items would need to be solved using an alternative strategy (e.g., simplification, equivalent component, cross-multiplication, see strategy coding section below). To balance the number of true and false items, we created three additional false pairs. See Appendix A for all 18 statements.

Strategy coding. We asked participants to report their strategy use immediately after each trial in the number line estimation and equivalence tasks. For the number line estimation task, we adopted coding schemes from past research (Sidney et al., 2018; Siegler & Thompson, 2014; Siegler et al., 2011). For any individual trial, a participant could report a strategy that we coded in multiple categories (e.g., segmenting and transformation). We also calculated the number of unique correct strategies a participant used (out of seven), since variable strategy use relates to estimation precision (Sidney et al., 2018).

We adapted coding schemes from research in fraction comparison (Fazio et al., 2016) to develop an initial coding scheme for the fraction equivalence task. We also included a number of new codes that were unique to the equivalence measure (e.g., componential transformation, simplification, common factor).

For both tasks, the first author and a research assistant coded each strategy independently for all participants, then the raters resolved disagreements through discussion until agreement was 100%. Initial agreement was 94.3% for the number line task (Cohen's Kappa = .75) and 95.8% for the equivalence task (Cohen's Kappa = .77) For the full list of strategies and examples, see Table 3 and 4.

Confidence ratings and familiarity. These tasks were the same as in Experiment 1 but not central to the present hypotheses and are not discussed further.

Demographic information. Participants reported the same demographic information in Experiment 2 as was reported in Experiment 1.

Table 3.

Number line estimation strategy codes and average use.

Strategy Category	Strategies	Definition	Example	Average Use
Transformation: *Transforming the fraction into a different number or form that is easier to estimate*	Rounding	Rounding: Approximating the fraction to a similar number.	*Approximated [54/63] as 5/6*	60.76%
	Simplifying	Simplifying: Transforming the fraction into a different fraction.	*I reduced [24/84] by 4 to 6/21.*	
	Translating	Translating: Transforming the fraction into a different form such as a percentage	*I know that 5/8ths is close to 0.6.*	
Segmentation*:* *Dividing the line into segments or parts.*	Halves reference	Any reference to half, the midpoint of the line, or segmenting the line into two parts.	*I divided the line in half then split the half line by 3.*	29.32%
	Denominator	Segmenting the line into parts based on the size of the denominator.	*Divide the line into 7ths and move the dot to 6.*	15.19%
	Other	Any other segmenting of the line.	*I divided the lower half of the line into three sections.*	3.32%
Magnitude: *Referencing the magnitude or size of the fraction.*	Magnitude	Placing a fraction based on the size.	*2/8 is 0.25 and 2/7 is a little bigger than that.*	33.02%
Spatial: *Referencing a spatial location or how close/far a value is to a reference point .*	Spatial	Placing a fraction based on a spatial location, such as further to the left or right.	*I went just past half.*	22.68%
	Landmarks	Placing a fraction based on how close or far it is to a landmark, such as 0, ½, or 1.	*6/7 is close to 1.*	
Visualization: *Thinking of the fraction using a visual representation other than the number line*	Visual	Placing the fraction based on how it is represented visually, such as parts of a pizza or pie.	*I thought about 7 pieces of pie and then taking away one piece.*	1.58%
Independent Components: *Referencing how big or small the components are independently*.	Components	Placing a fraction based solely on the size of the numerator and/or denominator.	*I put it closer to 0 because it being a single digit.*	1.27%

Note. Examples in italics are strategies reported by participants. Average use is calculated by dividing the total frequency each strategy was coded across participants by the total number of possible opportunities for the strategy to be reported.

Table 4.

Strategy Category	Strategies	Definition	Example	Average Use
Transformation: *Transforming the fraction into a different number or form that is easier to estimate*	Equivalence / Simplification*	Simplifying a fraction or transforming one of the fractions into a larger equivalent form.	*Divide the right by 9/9 and it becomes 4/9*	46.71%
	Common Components*	Transforming one or both fractions to have an equivalent numerator or denominator	*I reduced both sides so the denominators were 7 (divided the left side by 3 and the right side by 6)*	8.91%
	Componential	Attempting to see if the components of one fraction can evenly divide or multiply into the components of the other fraction.	*4 does not go into 35 cleanly therefore the statement is false.*	14.21%
	Vague	Strategies that suggest the participant transformed without a clear indication of how.	*I used the strategy of division to simplify*	12.53%
Cross Multiplication: *Comparing the whole number result of multiplying the cross components*	Cross Multiply*	Multiply the denominators by the numerators of the other fraction and comparing the whole number result.	*12(6)=72 and 72(1)=72.*	1.87%
Magnitude: *Referencing the magnitude or size of fractions.*	Magnitude or Half-reference*	Referencing the size of one or both fractions or indicating the fractions are on the opposite or same sides of one-half.	*54/63 is a lot more than 2/7*	10.72%
Division: *Numerator goes into the denominator same / different amount of times.*	Numerator-Denominator*	Divide each denominator by the numerator to see if the result is the same or different.	*both top numbers go into the bottom 6 times*	1.74%
Parity: *Referencing whether the components of fractions are odd or even.*	Parity	Deciding that fractions with odd or even components either can or cannot be equivalent.	*odd and even numbers not equal*	1.00%
Common factor: *Finding a least or greatest common factor.*	Common Factor	Noting that there is, or is not, a greatest or least common factor for the components of both fractions.	*common factor 9*	5.56%

Fraction equivalence task strategy codes and average use.

Note. Examples in italics are strategies reported by participants. Average use is calculated by dividing by the total frequency each strategy was coded across participants by the total number of possible opportunities for the strategy to be reported.

Procedure

Participants accepted the HIT through MTurk and were directed to the Qualtrics website to complete the survey on their own electronic device. Participants took an average of 37 minutes (range = 7 to 155min)[6] to complete the survey and were compensated $3.00. Participants were randomly assigned to complete the number line estimation task after (activated condition) or before (non-activated condition) completing the fraction equivalence task. After each trial in both the number line estimation and equivalence tasks, participants provided a written report of the strategy they used to solve the problem. At the end of the experiment, participants rated the familiarity of fractions and then answered demographic questions.

Results

Preliminary Analyses

We first confirmed random assignment was effective by examining differences in demographics and equivalence knowledge. There were no differences in age, $t(98) = 1.05$, $p = .298$, or equivalence scores, $t(98) = 0.58$, $p = .560$, between conditions. There was an approximately equal distribution of self-reported race, gender, and education level across conditions (all Chi-Square *p*s > .65). Estimation precision did not differ on the number line tasks between the two randomized orders presented within each condition (*t*s <1, *p*s >.6), so we collapsed number line order within each condition. As can be seen in Table 5, and replicating our results from Experiment 1, equivalence accuracy and PAE were related.

[6] Participants who took less than 10 minutes to complete the survey did not have codable strategy report data.

In the following sections, we report how components, condition, and equivalence knowledge affected adults' estimation performance. We examined estimation performance using the same analytic approach as used in Experiment 1. Then, we report the relation between strategy use and estimation performance. We also report the relation between strategy use and performance in the equivalence task. Finally, we report whether adults with more knowledge of equivalence are more accurate because they use more adaptive strategies.

Table 5.

Descriptive statistics and correlations between number line estimation and equivalence.

	N (α)	Mean (SD)	1	2
1. Number line estimation PAE	100 (.89)	7.5% (8.0%)	-	-
2. Equivalence Accuracy	100 (.72)	86.3% (13.6%)	**-.558******	-

Note. PAE = percent absolute error and is inversely related to accuracy. ** $p < .001$.

Number line estimation

Estimation precision. We predicted that adults' estimates of fractions with larger components would be less precise than estimates of equivalent fractions with smaller components but that equivalence knowledge would moderate this effect. We also explored whether activating knowledge of fraction equivalence reduced the effects of components or interacted with prior knowledge of equivalence. To test our hypotheses, we conducted a 2 (condition: activated vs. non-activated) x 2 (component size: small vs. large) between-within GLM with equivalence knowledge as a concomitant variable[7].

Within-subject effects revealed a significant effect of component size, $F(1, 96) = 6.29$, $p = .014$, $n_p^2 = .06$, and an interaction between component size and equivalence knowledge, F(1,

[7] As with Experiment 1, this analysis differs from our pre-registered difference score approach, but the conclusions are similar when analyzed using either approach. We chose to run a single GLM to reduce the number of necessary analyses to test our hypotheses.

96) = 5.84, p = .018, n_p^2 = .06. The interactions between component size and condition as well as between component size, condition, and equivalence knowledge were both non-significant (both ps > .6; see table B4 for full model).

In line with our first hypothesis (H1), follow-up planned comparisons revealed that the effect of components was only significant when equivalence knowledge was low. As can be seen in Figure 3, when modeled at low levels of equivalence knowledge, adults were less precise when they estimated fractions with large components (M = 13.0%, SE = 1.0%) relative to fractions with small components (M = 11.4%, SE = 1.1%), F(1, 96) = 5.22, p = .025, n_p^2 = .05. When modeled at average or high levels of equivalence knowledge, there were no differences in adults' estimation precision for fractions with small or large components (both ps > .2). We also confirmed that this pattern of results holds when we excluded participants who did not have codable strategy report data (described in the strategy report section). Indeed, the interaction between equivalence knowledge and component size remained significant (F[1, 75] = 5.24, p = .025, n_p^2 = .07), and the pairwise comparisons at higher and lower equivalence scores were similar to results when including the participants without codable strategy data (Table B6). Thus, as in Experiment 1, fraction equivalence knowledge was important for reducing the effects of component size on estimation precision.

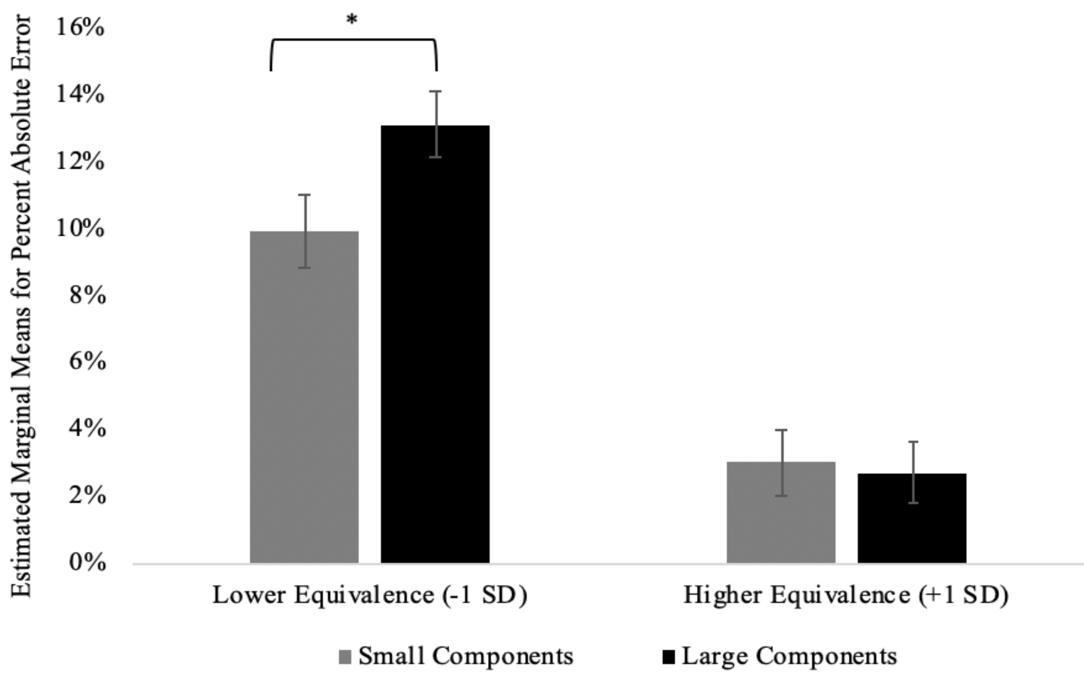

Figure 3. The effect of component size on PAE on estimated marginal means modeled when equivalence scores were higher (+ 1 SD) or lower (- 1 SD). *Note*. PAE is a measure of error and is inversely related to accuracy.

Even though we increased the size of components, adults with average levels of equivalence knowledge estimated small- and large-component fractions with similar precision. This surprising result may be due to reactivity from reporting strategies after each trial, which we discuss further in the discussion (e.g., Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Rittle-Johnson, Loehr, & Durkin, 2017). Indeed, the 39 participants who had at least one number line estimation trial with no codable strategy report data had significantly higher PAE (M = 12.14%, SD = 11.0%) than the 61 adults who reported codable strategy data on all trials (M = 4.5%, SD = 2.08%), $t(39.75) = 4.31$, $p < .001$, $d = .97$. Thus, providing quality strategy reports related to more precise number line estimates. However, the current experimental design cannot shed light on the direction of these effects; strategy reporting might have led to more precise estimates, or individuals who are more precise might have been able to provide codable strategy reports. We did replicate our main behavioral finding from Experiment 1: when equivalence knowledge was lower, adults were less precise when estimating large-component compared to small-component fractions.

Raw estimates. We ran a parallel GLM to examine the effects of component size and equivalence knowledge on raw estimates. However, there was no effect of components and no interaction with equivalence knowledge or condition (all *p*s > .18 See Table B5 for the full model).

However, we found that excluding participants without codable strategy data (described below) changed the results. We conducted a 2 (condition: activated vs. non-activated) x 2 (component size: small vs. large) between-within GLM with equivalence as a concomitant variable. The main effect of component size was non-significant, but component size interacted with condition, $F(1, 75) = 4.63$, p = .034, $n_p^2 = .06$, and this interaction was qualified by a marginal three way interaction with equivalence knowledge, $F(1, 75) = 3.57$, $p = .063$, $n_p^2 = .05$ (Table B7). The results of follow-up tests are consistent with our behavioral data from Experiment 1: when equivalence knowledge was lower, adults who had recently completed a fraction task fell back on componential heuristics, such as *smaller components = larger magnitudes*.

Strategy Use

Next, we examined participants' strategy use in the number line estimation and equivalence task. As pre-registered, we excluded participants who did not provide codable strategy reports on over half of the trials. This resulted in 21 participants being excluded from the number line task and 22 participants excluded from the equivalence task.

Strategy use and estimation precision. In line with overlapping waves theory, adults used a variety of different strategies to estimate fractions on number lines. Adults used between zero and seven different correct strategies across trials (*M* = 3.96, *SD* = 1.60). One participant guessed or used intuition on all trials, and only one participant used all seven correct strategies at least once across the 12 estimation trials. As can be seen in Table 6, and in line with our second hypothesis (H2a), adults who used magnitude strategies more frequently, had lower overall percent absolute error (PAE). Similarly, adults who used the transformation or segmentation strategy more frequently, or who used a variety of different strategies, had lower PAE.

Table 6.

Correlations between strategy use and performance on the number line estimation and equivalence task.

	N	Mean (SD)	1	2	3	4	5	6	7
1. *Number Line Magnitude PAE*[a]	79	5.7% (4.7%)	-	-	-	-	-	-	-
2. *Equivalence Accuracy*	79	87.9% (10.9%)	**-.430****	-	-	-	-	-	-
3. *Transformation Strategy*	79	7.29 (3.8)	**-.344****	**.508****	-	-	-	-	-
4. *Magnitude Strategy*	79	3.96 (3.60)	**-.276***	.122	.165	-	-	-	-
5. *Any Segmentation Strategy*	79	5.86 (4.12)	**-.311****	-.066	**-.294***	.159	-	-	-
6. *Spatial Strategy*	79	2.72 (3.14)	.205	-.072	-.132	.064	.179	-	-
7. *Visualization Strategy*	79	0.19 (0.86)	.050	**-.274***	-.152	.027	.000	.124	-
8. *Strategy Variability (out of 7)*	79	3.96 (1.6)	**-.302****	-.087	-.185	**.410****	**.677****	**.306***	.209

Note. Differences in correlation values from Table 5 are due to filtering out participants who did not provide codable strategy reports. ** $p < .01$, * $p < .05$. [a]PAE = percent absolute error and is inversely related to estimation precision.

Fraction equivalence knowledge and adaptive strategy use. To examine whether strategy use explained the relation between equivalence knowledge and estimation precision, we explored several indirect effect models. In all models, we controlled for condition and used bootstrapping to estimate coefficients and confidence intervals using the PROCESS macro for SPSS (Hayes, 2012). First, we examined whether adults with more equivalence knowledge had lower PAE indirectly through their strategy variability. However, there was no indirect effect of equivalence knowledge on PAE through strategy variability (b = .01, 95% CI [-.02, .05]). Next, we examined whether equivalence knowledge related to estimation precision through the use of *adaptive* strategies. Two strategy types might be considered particularly adaptive in our experiment. First, magnitude strategies are adaptive on *any* trial during number line estimation since correct use of these strategies will always lead to precise estimates. This strategy may be more difficult when estimating fractions with large components. When estimating large-component fractions, adults may need to use the second adaptive strategy--transformation--prior to estimating the fraction magnitude. Use of transformation strategies often *indirectly* indicates magnitude understanding. For example, an adult reported that they reduced 24/84 "by 4 to 6/21." Maintaining the correct ratio between components demonstrates accurate magnitude understanding, and this indirect reference to magnitude has been coded as adaptive in past research (Sidney et al., 2018).

First, we tested whether equivalence knowledge indirectly related to more precise estimates overall through the use of magnitude strategies. However, this model was non-significant, indirect effect: b = -.01, 95% CI [-.04, .01]. Next, we tested whether there was an indirect effect of adults' equivalence knowledge on PAE for large component trials through use of transformation strategies on those trials. We examined transformation strategies and PAE on large-component trials because these are the trials when adults' were biased. As can be seen in Figure 4, higher equivalence knowledge related to a higher frequency of using the transformation strategy on large-component trials, $b = 9.47$, $t(76) = 5.76$, $p < .001$, 95% CI [6.19, 12.74], use of this strategy related to more precise estimates on large-component trials, $b = -0.01$, $t(75) = 2.93$, $p = .004$, 95% CI [-0.01, -0.002], and the indirect effect on large-component PAE was also significant, $b = -0.08$, 95% CI [-0.18, -0.01]. This was not the case when examining the role of transformation strategies overall on average PAE. In other words, equivalence knowledge only had an indirect effect on PAE through the frequency of transforming *large-component* fractions, but not through transformation on all trials.

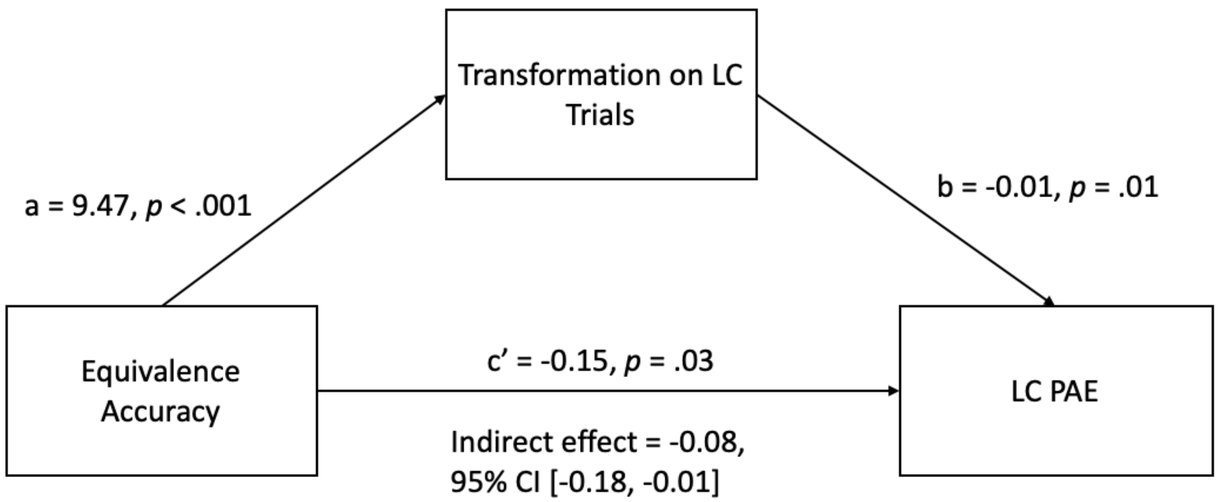

Figure 4. Indirect effect of equivalence knowledge on percent absolute error for large component trials through the frequency of using the transformation strategy when estimating large component fractions. *Note*. PAE is inversely related to precision. LC = large component.

Equivalence task strategy use. As with number line estimation, we analyzed adults' optimal strategy variability (out of 5 possibly strategies). For the equivalence task, we coded simplification, common component, cross-multiplication, magnitude, and division strategies as optimal, since each of these strategies would lead to the correct response if executed correctly. Adults used between 0 and 5 optimal strategies on the equivalence task ($M = 2.40$, $SD = 1.05$).

Five participants did not use an optimal strategy on any trial, and only one used all five optimal strategies at least once.

We predicted that use of magnitude strategies in the equivalence task would relate to higher accuracy. In line with our second hypothesis (H2b), adults who used the magnitude strategy more frequently had higher performance on the equivalence task, $r = .311$, $p = .004$. Higher performance on the equivalence task also related to the frequency adults used simplification, $r = .677$, $p < .001$, and common component, $r = .344$, $p = .001$, strategies. However, cross-multiplication and division--the other two strategies we coded as adaptive--were unrelated to performance, possibly because they were used less frequently than the other adaptive strategies overall (1.85% and 1.74%, see Table 4). We also explored whether strategy variability in the equivalence task related to overall performance. In line with overlapping waves theory (Siegler, 1996, 2005, 2016b), variability among optimal strategies related to higher accuracy on the equivalence task, $r = .318$, $p = .003$. However, variability among all strategies was unrelated to accuracy.

Experiment 2 Discussion

In Experiment 2, we examined how components, individual differences in equivalence knowledge, and recent experiences with fraction equivalence impacted adults' reasoning about fraction magnitudes. In line with our first hypothesis (H1), adults with less knowledge of fraction equivalence were less precise when estimating fractions with larger compared to smaller components. However, adults with average or above average levels of equivalence knowledge estimated equivalent fractions with small and large components with similar precision. Thus, we replicated our results from the first experiment: whole number components impacted adults' ability to precisely estimate fraction magnitudes, especially when equivalence knowledge was low. Importantly, we examined the strategies adults used to reason about fraction magnitudes.

Adults used a variety of different strategies when they estimated fractions on number lines and evaluated whether two fractions were equivalent. In line with our second hypothesis (H2a-b), adults who more frequently used strategies based on the holistic magnitude of fractions had better performance in both tasks. Importantly, individual differences and contextual factors influenced the strategies that adults used in the number line estimation task. In the estimation task, adults with more equivalence knowledge were more likely to use the transformation strategy when estimating large component fractions, and use of this strategy related to lower overall percent absolute error (PAE). Thus, we extended findings from the first experiment: individual differences in knowledge of fraction equivalence related to more precise number line estimates, in part due to adaptive strategy use.

General Discussion

Adults' performance in fraction magnitude tasks sometimes reflects whole number bias, but other times reflects accurate reasoning. Across two experiments, contextual factors and individual differences influenced adults' fraction reasoning. Adults demonstrated whole number bias when it was difficult to reason about fractions as holistic magnitudes: when estimating under a time constraint, when estimating fractions with large components, or when comparing fractions close in decimal distance. We extend past research by demonstrating that individual differences in knowledge of fraction equivalence were more important than inhibition or updating abilities for reducing whole number bias during number line estimation. When equivalence knowledge was higher[8], adults were not biased by fraction components when estimating fractions on number lines.

These novel results advance research on adults' fraction reasoning in several important ways. First, we replicated research demonstrating that adults *can* reason about fractions as holistic magnitudes (e.g., Gabriel et al., 2013; Obersteiner et al., 2012; Schneider & Siegler, 2010; Sprute & Temple, 2011; Zhang et al., 2014), even though some adults still demonstrate whole number bias. Second, to our knowledge, we are the first to demonstrate fraction equivalence knowledge moderates whole number bias in adults. We linked fraction equivalence knowledge to adaptive strategy use, thereby extending past research and theories on strategy variability and whole number bias (e.g., Alibali & Sidney, 2015; Fazio et al., 2016; Sidney et al., 2018; Siegler, 1996, 2005, 2016b; Siegler & Thompson, 2014; Siegler et al., 2011). Establishing a relation between equivalence knowledge and fraction reasoning elucidates for *whom* components negatively impact fraction reasoning, and linking equivalence knowledge to adaptive strategy use elucidates *how* adults adaptively use strategies to reason about fractions.

Theoretical Implications

Strategic variability (Siegler, 1996, 2005, 2016b) occurs in many domains: arithmetic (e.g., LeFevre, Bisanz, et al., 1996; LeFevre et al., 1996), memory (e.g., Schneider, Kron-Sperl, & Hünnerkopf, 2009), locomotion (e.g., Adolph, Vereijken, & Denny, 1998), performance on false belief problems (Flynn, O'Malley, & Wood, 2004), and even when playing tic-tac-toe (Crowley & Siegler, 1993). Similar to strategic variability in other domains, the results from our current experiments indicated that there was systematic variability in *how*, *when*, and for *whom* whole number components impacted adults' fraction reasoning (see Alibali & Sidney, 2015).

[8] We use a person-centered interpretation in the discussion of our results, Note that we modeled these findings at higher and lower levels of equivalence scores.

Our findings build on past work demonstrating that the effects of whole number components on adults' fraction reasoning vary *within* individuals who are solving problems in different contexts and *between* individuals who possess different prior knowledge (DeWolf & Vosniadou, 2015; Fazio et al., 2016; Obersteiner et al., 2013; Schneider & Siegler, 2010; Sidney et al., 2018; Zhang et al., 2016). For example, even expert mathematicians draw on their knowledge of whole numbers to compare fractions when all fraction pairs in the task share a common component, yet use fraction magnitude strategies when components differ from fraction-to-fraction (Obersteiner et al., 2013). Experts were faster to compare fractions with common components when the correct answer was consistent with whole number knowledge (e.g., 6/11 < 8/11 and 6 < 8) relative to when the correct answer was inconsistent with whole number knowledge (e.g., 6/13 > 6/47 even though 47 > 13). Experts had to inhibit their whole number strategies when comparing fractions with common numerators, but did not have to do so when comparing fractions with common denominators or when making comparisons that elicited fraction magnitude strategies (i.e., fraction pairs with no common components). Thus, in some contexts, whole number strategies are adaptive; in others, these same strategies result in errors or increased response times. These findings are consistent with other work that demonstrates adults adapt their strategy selection based on characteristics of the problem (e.g., Fazio et al., 2016; Schneider & Siegler, 2010).

In the context of estimating one fraction at a time in our experiments, whole number magnitudes may not need to be inhibited. Indeed, adults' ability to inhibit whole number magnitude knowledge in Experiment 1 only marginally moderated the effect of whole number components in the estimation task, but related to overall comparison performance. In contrast, updating abilities related to overall estimation *and* comparison performance. Thus, inhibition of whole number knowledge may be more important in the context of comparing fraction magnitudes than in the context of estimating because the comparison fraction is visible on the screen, and therefore the direct comparison between numerator and denominator components must be inhibited. Alternatively, the correlation between performance on the fraction comparison and inhibition tasks might be due to both tasks involving *comparisons* between numbers rather than due to inhibition of magnitude knowledge generally (i.e., shared task variance). Future work might test this possibility by including domain-specific and domain-general measures of inhibition.

The role of different mental processes (e.g., inhibition and updating) in reasoning about fractions in different task contexts does not conflict with conceptual change approaches to whole number bias. For example, our data do not rule out the possibility that whole numbers might be processed differently than fractions. However, our data point to the fact that people often integrate whole number and fraction knowledge as they adaptively choose which strategy to employ from trial-to-trial. For example, consider how an adult with more knowledge of fraction equivalence might estimate the fraction 12/72 on a number line. This person's understanding of fraction equivalence might be activated when they see 12/72; this fraction is difficult to estimate in its current form, but it can be transformed. To transform the fraction, the person might draw on their knowledge of whole number multiplication or division. Thus, with relevant prior knowledge of fraction equivalence, the appropriate *combination* of fraction and whole number knowledge leads to adaptive strategy selection and execution.

Strategic variability accounts also suggest that relevant prior knowledge should be related to more precise estimates of fraction magnitudes (Alibali & Sidney, 2015; Schneider & Siegler, 2010; Siegler, 1996, 2005, 2016b). In our studies, we found that the relation between adults' number line estimation precision and their equivalence knowledge provided a compelling case for the relevance of prior knowledge. Further, we replicated the Bayes Factors models reported by Braithwaite and Siegler (2017) in which, for each participant, we compared models that predicted fraction estimates from fraction magnitudes (magnitude model), numerator and denominator components (componential model), or fraction magnitude *and* components together (hybrid model). For each model and participant, we calculated Bayes Factors relative to an intercept only model using the BayesFactor package for *R* (Morey, Rouder, & Jamil, 2015) and transformed the Bayes Factors into posterior probabilities (see Supplemental Analyses File). Across both experiments, equivalence knowledge related to having a higher posterior probability of being fit by a magnitude model, and a lower probability of being fit by a componential or hybrid model. It is currently unclear whether equivalence knowledge causes more precise magnitude representations (or vice versa) and whether effectively executing strategies causes more precise magnitude representations (or vice versa). However, the interrelations between equivalence knowledge, strategy use, contextual factors, and performance suggest that some experiences may be more likely to support adaptive strategy selection and execution than others.

Environmental Experiences and Educational Implications

Why are some participants more likely to possess a holistic representation of magnitudes than others? Holistic magnitude representations may require particular environmental experiences (e.g., with fraction equivalence) that build on people's intuitive ability to directly reason about ratios, such as fractions (e.g., Jacob & Nieder, 2009; Binzak, Matthews, & Hubbard, 2019; Lewis, Matthews, & Hubbard, 2015; Matthews & Chesney, 2015; Matthews & Ellis, 2018; Matthews & Lewis, 2017). New research suggests that people possess a ratio processing system (RPS; Matthews, Lewis, & Hubbard, 2015) that can be used to directly and automatically process ratios. The RPS may be leveraged to support learning about fraction magnitudes through the frequency with which fractions are encountered in the environment. For example, fractions like 1/4, 1/2, or 3/4 are encountered more frequently in the environment than other fractions (e.g, Braithwaite & Siegler, 2018), and adults are more familiar with, and confident when estimating common compared to uncommon fractions (Fitzsimmons, Thompson, & Sidney, 2019). Furthermore, adults quickly compare common fractions as holistic magnitudes, yet they strategically compute the magnitude of less common fractions (Liu, 2017). Our Bayes Factor analysis provided evidence that more extensive experience with fraction equivalence may be one way to facilitate processing fractions based on holistic magnitudes.

The role of equivalence knowledge and strategy use in adults' whole number bias provides a foundation for researchers to develop evidence-based instructional sequences to facilitate fraction learning (e.g., Moss & Case, 1999; Tian & Siegler, 2018). When learning about equivalent ratios in sixth grade, some children shift from componential-based reasoning about fractions towards accurate magnitude representations (Common Core State Standards, CCSS-M, 6.RP.A.3, 2010; Braithwaite & Siegler, 2017; Gabriel et al., 2013; Van Hoof, Degrande, Ceulemans, Verschaffel, & Van Dooren 2018; Van Hoof et al., 2015). Children's additional experience with equivalence may help them recognize the importance of the relation between components for accurate reasoning about fraction magnitudes. However, the relation between equivalence knowledge and fraction magnitude understanding is likely bi-directional because evidence suggests conceptual and procedural knowledge (Rittle-Johnson, 2017; Rittle-Johnson, Schneider, & Star, 2015; Rittle-Johnson, Siegler, & Alibali, 2001) and fraction magnitude and arithmetic knowledge develop iteratively (Bailey, Hansen, & Jordan, 2017). Future research could examine whether improving equivalence knowledge subsequently improves children's use of magnitude-based strategies and diminishes the effects of whole number bias, or vice versa.

To improve fraction magnitude understanding, future researchers might test an intervention using self-explanation prompts. In Experiment 2, adults who provided higher quality strategy reports had more precise number line estimates than those who could not, consistent with past work suggesting that self-explanation prompts can improve performance (Chi et al., 1989; Chi, Leeuw, Chiu, & Lavancher, 1994; DeCaro, Rotar, Kendra, & Beilock, 2010; McEldoon, Durkin, & Rittle-Johnson, 2013; Rittle-Johnson, Fyfe, Loehr, & Miller, 2015; Rittle-Johnson et al., 2017; Rittle-Johnson & Loehr, 2017; VanLehn, Jones, & Chi, 1992). Self-explanations are thought to help students address misconceptions and integrate knowledge by directing attention to important structural, rather than surface, features of problems (e.g. Rittle-Johnson et al., 2017). Misconceptions and attention to whole number components--surface features of fractions--are the most likely causes of whole number bias during fraction magnitude reasoning. Thus, self-explanation prompts might be ideal for improving fraction magnitude and equivalence understanding by directing participants' attention to holistic magnitudes rather than to numerator and denominator components.

What Makes Fraction Equivalence Knowledge Important?

Understanding fraction equivalence can be considered a highly specialized form of magnitude knowledge: it requires understanding a *precise*, rather than approximate, relation between fractions. That is, when participants decide that two fractions are equivalent, they presumably know the exact magnitude of each fraction (1/2 and 15/30 both = 0.5). When they decide which of two fractions is larger, they can approximate the magnitude of the two fractions (i.e., fraction A is greater than 1/2 and fraction B is less than 1/2, so fraction A is largest), and this will lead to the correct answer except for situations in which it is too difficult to differentiate approximated magnitudes, such as for fractions that are very close in decimal distance.

What leads to the development of fraction equivalence precision? Practice with whole number computational skills, such as multiplication and division, likely contribute to the development of this ability. Whole number knowledge can *help*, rather than hinder, fraction learning (e.g., Namkung, Fuchs, & Koziol, 2018; Sidney & Alibali, 2015, 2017; Sidney, Thompson, & Rivera, 2019; Siegler et al., 2011). For example, children's proficiency with multiplication and division predicts how well they will learn fraction concepts (Namkung et al., 2018), and division proficiency predicts later algebra performance (Siegler et al., 2012). Furthermore, activating students' knowledge of whole number division facilitates their ability to represent fraction division problems (e.g., Sidney & Alibali, 2015, 2017). Multiplication or division is often necessary to execute fraction transformation strategies, the most commonly reported strategy on the fraction equivalence task (e.g., simplify 36/81 by dividing by 9/9; see Table 4). Thus, our measure of equivalence knowledge likely captures some of adults' proficiency with whole number multiplicative reasoning (Steffe & Olive, 2009). Future work might explore whether knowledge of fraction equivalence mediates the relation between earlier whole number knowledge and later fraction understanding.

Conclusion

Children and adults often inappropriately apply their whole number knowledge while reasoning about fractions (Alibali & Sidney, 2015; Braithwaite & Siegler, 2017; DeWolf & Vosniadou, 2015; Ni & Zhou, 2005; Siegler & Lortie-Forgues, 2015; Siegler et al., 2011; Vamvakoussi, 2015). We discovered that individual differences in fraction equivalence knowledge moderate how whole number components impact adults' fraction reasoning. We are the first to demonstrate that adults with more equivalence knowledge can access more precise representations of fraction magnitudes, in part, due to their adaptive use of transformation strategies. Adapting strategies based on knowledge of fraction equivalence and contextual factors provides evidence that strategic variability accounts of cognition can explain adults' fraction reasoning and whole number bias errors. The role of equivalence knowledge and transformation strategies in adults' fraction reasoning provides an important foundation for researchers to explore the link between whole number and fraction knowledge, design interventions that aim to reduce whole number bias earlier in development, and improve fraction magnitude understanding.

References

Adolph, K. E., Vereijken, B., Denny, M. A. (1998). Learning to crawl. *Child Development*, *69*(5), 1299-1312. doi:10.1111/j.1467-8624.1998.tb06213.x.

Alibali, M. W., & Sidney, P. G. (2015). Variability in the natural number bias: Who, when, how, and why. *Learning and Instruction, 37*, 56-61.

Bailey, D. H., Hansen, N., & Jordan, N. C. (2017). The codevelopment of children's fraction arithmetic skill and fraction magnitude understanding. *Journal of Educational Psychology*, *109*(4), 509.

Behr, M. J., Wachsmuth, I., Post, T. R., & Lesh, R. (1984). Order and equivalence of rational numbers: A clinical teaching experiment. *Journal for Research in Mathematics Education*, 323-341.

Binzak, J., Matthews, P. G., & Hubbard, E. M. (2019). On Common Ground: Evidence for an Association Between Fractions and the Ratios They Represent. *Retrieved from* https://psyarxiv.com/jxhtv/.

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*(6), 1410.

Booth, J. L., & Newton, K. J. (2012). Fractions: Could they really be the gatekeeper's doorman? *Contemporary Educational Psychology*, *37*(4), 247-253.

Braithwaite, D. W., & Siegler, R. S. (2016). Numerical Development. *Annual Review of Psychology, 68,* 187-213. doi:10.1146/annurev-psych-010416-044101.

Braithwaite, D. W., & Siegler, R. S. (2017). Developmental changes in the whole number bias. *Developmental Science*, *21*(2).

Braithwaite, D. W., & Siegler, R. S. (2018). Children learn spurious associations in their math textbooks: Examples from fraction arithmetic. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Campbell, J. I., & Alberts, N. M. (2009). Operation-specific effects of numerical surface form on arithmetic strategy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 999.

Campbell, J. I., & Austin, S. (2002). Effects of response time deadlines on adults' strategy choices for simple addition. *Memory & Cognition*, *30*(6), 988-994.

Caviola, S., Carey, E., Mammarella, I. C., & Szücs, D. (2017). Stress, time pressure, strategy selection and math anxiety in mathematics: A review of the literature. *Frontiers in Psychology*, *8*, 1488.

Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145-182.

Chi, M. T., De Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*(3), 439-477.

Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education*, *3*(2), 63-68.

Crescenzi, A., Kelly, D., & Azzopardi, L. (2015, August). Time pressure and system delays in information search. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 767-770). ACM.

Cronbach, L. J. & Furby, L. (1970). How we should measure "change" - or should we? *Psychological Bulletin*, *74*(1), 68-80.

Crooks, N. M., & Alibali, M. W. (2013). Noticing relevant problem features: Activating prior knowledge affects problem solving by guiding encoding. *Frontiers in Psychology*, *4*, 884.

Crowley, K. & Siegler, R. S. (1993). Flexible strategy use in young children's tic-tac-toe. *Cognitive Science*, 17(4), 531-561. doi:10.1207/s15516709cog1704_3.

DeCaro, M. S. (2016). Inducing mental set constrains procedural flexibility and conceptual understanding in mathematics. *Memory & Cognition*, *44*(7), 1138-1148.

DeCaro, M. S., Rotar, K. E., Kendra, M. S., & Beilock, S. L. (2010). Diagnosing and alleviating the impact of performance pressure on mathematical problem solving. *Quarterly Journal of Experimental Psychology*, *63*(8), 1619-1630.

Dadon, G., & Henik, A. (2017). Adjustment of control in the numerical Stroop task. *Memory & Cognition*, *45*(6), 891-902.

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

Dehaene, S. (2003). The neural basis of the Weber–Fechner law: a logarithmic mental number line. *Trends in Cognitive Sciences*, *7*(4), 145-147.

Dehaene, S. (2011). *The number sense: How the mind creates mathematics*. New York, NY: Oxford University Press.

Dumas, J. E., Johnson, M., & Lynch, A. M. (2002). Likableness, familiarity, and frequency of 844 person-descriptive words. *Personality and Individual Differences*, *32*(3), 523-531.

DeWolf, M., & Vosniadou, S. (2015). The representation of fraction magnitudes and the whole number bias reconsidered. *Learning and Instruction*, *37*, 39-49.

Edwards, J. R. (1995). Alternatives to difference scores as dependent variables in the study of congruence in organizational research. *Organizational Behavior and Human Decision Processes*, 64, 307-324.

Edwards, J. R. (2001). Ten difference score myths. *Organizational Research Methods*, 4(3), 265-287.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. The MIT Press.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). The MIT Press.

Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, and Activity*, *5*(3), 178-186.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160.

Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology*, *123*, 53-72.

Fazio, L. K., DeWolf, M., & Siegler, R. S. (2016). Strategy use and strategy choice in fraction magnitude comparison. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(1), 1.

Fitzsimmons, C. J., Thompson, C. A., & Sidney, P. G. (2019). Confident or familiar? The role of familiarity ratings in adults' confidence judgments when estimating fraction magnitudes. Proceedings of the 41st Annual Psychology of Mathematics Education-North America Conference. St. Louis, MO.

Fitzsimmons, C. J., Thompson, C. A., & Sidney, P. G. (2019). Do adults treat equivalent fractions equally? Adults' strategies and errors in fraction reasoning. *Online repository of pre-registrations, data, and syntax to be made public upon acceptance for publication*. Retrieved from: *https://osf.io/4uygd/*. doi: 10.17605/OSF.IO/4UYGD

Flynn, E., O'Malley, C., & Wood, D. (2004). A longitudinal, microgenetic study of the emergence of false belief understanding and inhibition skills. *Developmental Science*, 7(1), 103-115. doi:10.1111/j.1467-7687.2004.00326.x

Gabriel, F. C., Szücs, D., & Content, A. (2013). The development of the mental representations of the magnitude of fractions. *PloS One*, *8*(11),

Gillard, E., Van Dooren, W., Schaeken, W., & Verschaffel, L. (2009). Dual processes in the psychology of mathematics education and cognitive psychology. *Human Development*, *52*(2), http://dx.doi.org/10.1159/000202728.

Gómez, D. M., Jiménez, A., Bobadilla, R., Reyes, C., & Dartnell, P. (2015). The effect of inhibitory control on general mathematics achievement and fraction comparison in middle school children. *ZDM*, *47*(5), 801-811.

Handel, M. J. (2016). What do people do at work? *Journal for Labour Market Research*, *49*(2), 177-197.

Hayes, A. F. (2012). Process SPSS macro. *Computer Software and Manual.* Available from http://www afhayes.com/public/process.pdf.

Jacob, S. N., & Nieder, A. (2009). Notation-independent representation of fractions in the human parietal cortex. *Journal of Neuroscience*, *29*(14), 4652-4657.

Judd, C. M., Kenny, D. A., & McClelland, G. H. (2001). Estimating and testing mediation and moderation in within-subject designs. *Psychological Methods*, *6*(2), 115.

Kallai, A. Y., & Tzelgov, J. (2009). A generalized fraction: an entity smaller than one on the mental number line. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1845.

Kallai, A. Y., & Tzelgov, J. (2012). When meaningful components interrupt the processing of the whole: The case of fractions. *Acta Psychologica*, *139*(2), 358-369.

Kaufmann, L., Ischebeck, A., Weiss, E., Koppelstaetter, F., Siedentopf, C., Vogel, S. E., ... & Wood, G. (2008). An fMRI study of the numerical Stroop task in individuals with and without minimal cognitive impairment. *Cortex*, *44*(9), 1248-1255.

Kellogg, J. S., Hopko, D. R., & Ashcraft, M. H. (1999). The effects of time pressure on arithmetic performance. *Journal of Anxiety Disorders*, *13*(6), 591-600.

Kolkman, M. E., Hoijtink, H. J., Kroesbergen, E. H., & Leseman, P. P. (2013). The role of executive functions in numerical magnitude skills. *Learning and Individual Differences*, *24*, 145-151.

Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Development*, *78*(6), 1723-1743.

LeFevre, J. -A., Bisanz, J., Daley, K. E., Buffone, L., Greenham, S. L., & Sadesky, G. S. (1996). Multiple routes to solution of single-digit multiplication problems. *Journal of Experimental Psychology: General*, *125*(3), 284-306. Doi:10.1037/0096-3445.125.3.284.

LeFevre, J. -A., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in metal addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(1), 216-230.

Leibovich, T., Diesendruck, L., Rubinsten, O., & Henik, A. (2013). The importance of being relevant: Modulation of magnitude representations. *Frontiers in Psychology, 4*, 369.

Lewis, M. R., Matthews, P. G., & Hubbard, E. M. (2015). Neurocognitive architectures and the nonsymbolic foundations of fractions understanding. In *Development of Mathematical Cognition* (pp. 141-164). Academic Press.

Liu, F. (2017). Mental representation of fractions: It all depends on whether they are common or uncommon. *The Quarterly Journal of Experimental Psychology*, (just-accepted), 1-38.

Matthews, P. G., & Chesney, D. L. (2015). Fractions as percepts? Exploring cross-format distance effects for fractional magnitudes. *Cognitive Psychology*, *78*, 28-56.

Matthews, P. G., & Ellis, A. B. (2018). Natural alternatives to natural number: The case of ratio. *Journal of Numerical Cognition*, *4*(1), 19-58. doi:10.5964/jnc.v4i1.97.

Matthews, P. G., & Lewis, M. R. (2017). Fractions we cannot ignore: The nonsymbolic ratio congruity effect. *Cognitive Science*, *41*(6), 1656-1674.

Matthews, P. G., Lewis, M. R., & Hubbard, E. M. (2015). Individual differences in nonsymbolic ratio processing predict symbolic math performance. *Psychological Science*, *27*(2), 191-202. doi:10.1177/0956797615617799.

Maxwell, S. E. & Delaney, H. D., (2004). Designing experiments and analyzing data. Mahwah, NJ: Lawrence Erlbaum Associates.

McEldoon, K. L., Durkin, K. L., & Rittle- Johnson, B. (2013). Is self- explanation worth the time? A comparison to additional practice. *British Journal of Educational Psychology*, *83*(4), 615-632.

McNeil, N. M., & Alibali, M. W. (2005). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development*, *76*(4), 883-899.

Meert, G., Grégoire, J., & Noël, M. P. (2009). Rational numbers: Componential versus holistic representation of fractions in a magnitude comparison task. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1598-1616.

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*(3), 284-292.

Meert, G., Grégoire, J., Seron, X., & Noël, M. P. (2012). The mental representation of the magnitude of symbolic and nonsymbolic ratios in adults. *The Quarterly Journal of Experimental Psychology*, *65*(4), 702-724.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100.

Morey, R. D., Rouder, J. N., & Jamil, T. (2015). BayesFactor: Computation of Bayes factors for common designs. R package version 0.9. 12-2.

Moss, J., & Case, R. (1999). Developing children's understanding of the rational numbers: A new model and an experimental curriculum. *Journal for Research in Mathematics Education*, 122-147.

Namkung, J. M., Fuchs, L. S., & Koziol, N. (2018). Does initial learning about the meaning of fractions present similar challenges for students with and without adequate whole-number skill? *Learning and Individual Differences, 61*, 151-157.

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

Ni, Y. (2001). Semantic domains of rational numbers and the acquisition of fraction equivalence. *Contemporary Educational Psychology*, *26*(3), 400-417.

Ni, Y., & Zhou, Y. D. (2005). Teaching and learning fraction and rational numbers: The origins and implications of whole number bias. *Educational Psychologist*, *40*(1), 27-52.

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.

Opfer, J. E., & DeVries, J. M. (2008). Representational change and magnitude estimation: Why young children can make more accurate salary comparisons than adults. *Cognition*, *108*(3), 843-849.

Opfer, J. E., & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology*, *55*(3), 169-195.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science*, *17*(5), 407-413.

Revelle, W. (2018). psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych

Rittle- Johnson, B. (2017). Developing mathematics knowledge. *Child Development Perspectives*, *11*(3), 184-190.

Rittle-Johnson, B., Fyfe, E. R., Loehr, A. M., & Miller, M. R. (2015). Beyond numeracy in preschool: Adding patterns to the equation. *Early Childhood Research Quarterly*, *31*, 101-112.

Rittle-Johnson, B., & Loehr, A. M. (2017). Eliciting explanations: Constraints on when self-explanation aids learning. *Psychonomic Bulletin & Review*, *24*(5), 1501-1510.

Rittle-Johnson, B., Loehr, A. M., & Durkin, K. (2017). Promoting self-explanation to improve mathematics learning: A meta-analysis and instructional design principles. *ZDM*, *49*(4), 599-611.

Rittle-Johnson, B., Schneider, M., & Star, J. R. (2015). Not a one-way street: Bidirectional relations between procedural and conceptual knowledge of mathematics. *Educational Psychology Review*, *27*(4), 587-597.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93*(2), 346.

Schneider, W., Kron-Sperl, V., & Hünnerkopf, M. (2009). The development of young children's memory strategies: Evidence from the Würzburg longitudinal memory study. *European Journal of Developmental Psychology*, 6(1), 70-99. doi.org/10.1080/17405620701336802.

Schneider, M., Merz, S., Stricker, J., De Smedt, B., Torbeyns, J., Verschaffel, L., & Luwel, K. (2018). Associations of number line estimation with mathematical competence: A meta- analysis. *Child Development*, *89*(5), 1467-1484.

Schneider, M., & Siegler, R. S. (2010). Representations of the magnitudes of fractions. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1227.

Sidney, P. G., & Alibali, M. W. (2015). Making connections in math: Activating a prior knowledge analogue matters for learning. *Journal of Cognition and Development*, *16*(1), 160-185.

Sidney, P. G., & Alibali, M. W. (2017). Creating a context for learning: Activating children's whole number knowledge prepares them to understand fraction division. *Journal of Numerical Cognition*, *3*(1), 31-57, https://doi.org/10.5964/jnc.v3i1.71.

Sidney, P. G., Thalluri, R., Buerke, M. L., & Thompson, C. A. (2018). Who uses more strategies? Linking mathematics anxiety to adults' strategy variability and performance on fraction magnitude tasks. *Thinking & Reasoning*, *25*(1), 94-131.

Sidney, P. G., Thompson, C. A., & Rivera, F. D. (2019). Number lines, but not area models, support children's accuracy and conceptual models of fraction division. *Contemporary Educational Psychology*, *58*, 288-298.

Siegler, R. S. (1996). *Emerging Minds: The Process of Change in Children's Thinking*. New York, NY: Oxford University Press.

Siegler, R. S. (2005). Children's learning. *American Psychologist*, *60*(8), 769.

Siegler, R. S. (2016a). Magnitude knowledge: The common core of numerical development. *Developmental Science*, *19*(3), 341-361.

Siegler, R. S. (2016b). Continuity and change in the field of cognitive development and in the perspectives of one cognitive developmentalist. *Child Development Perspectives*, *10*(2), 128-133.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*(2), 428-444.

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., ... & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, *23*(7), 691-697.

Siegler, R. S., & Lortie- Forgues, H. (2014). An integrative theory of numerical development. *Child Development Perspectives*, *8*(3), 144-150.

Siegler, R. S., & Pyke, A. A. (2013). Developmental and individual differences in understanding of fractions. *Developmental psychology*, *49*(10), 1994.

Siegler, R. S., & Thompson, C. A. (2014). Numerical landmarks are useful—except when they're not. *Journal of Experimental Child Psychology*, *120*, 39-58.

Siegler, R. S., Thompson, C. A., & Opfer, J. E. (2009). The logarithmic- to- linear shift: One learning sequence, many tasks, many time scales. *Mind, Brain, and Education*, *3*(3), 143-150.

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology*, *62*(4), 273-296.

Smagorinsky, P. (1998). Thinking and speech and protocol analysis. *Mind, Culture, and Activity*, *5*(3), 157-177.

Smith, E. R, & Miller, F. D. (1978). Limits on perception of cognitive processes: A reply to Nisbett and Wilson. *Psychological Review, 85*(4), 355-362.

Sprute, L., & Temple, E. (2011). Representations of fractions: Evidence for accessing the whole magnitude in adults. *Mind, Brain, and Education, 5*(1), 42-47.

Stafylidou, S., & Vosniadou, S. (2004). The development of students' understanding of the numerical value of fractions. *Learning and Instruction*, *14*(5), 503-518.

Steffe, L. P., & Olive, J. (2009). *Children's fractional knowledge*. New York, NY: Springer Science & Business Media.

Tian, J., & Siegler, R. S. (2018). Which type of rational numbers should students learn first? *Educational Psychology Review*, *30*(2), 351-372.

Thevenot, C., Castel, C., Fanget, M., & Fayol, M. (2010). Mental subtraction in high- and lower skilled arithmetic problem solvers: Verbal report versus operand-recognition paradigms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), 1242-1255. doi:10.1037/a0020447

Thompson, C. A., & Opfer, J. E. (2010). How 15 hundred is like 15 cherries: Effect of progressive alignment on representational changes in numerical cognition. *Child Development*, *81*(6), 1768-1786.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. National Assessment of Educational Progress (NAEP), 1981. Mathematics Assessment. 1996.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. National Assessment of Educational Progress (NAEP), 2008. Mathematics Assessment. 2008.

Vamvakoussi, X. (2015). The development of rational number knowledge: Old topic, new insights. *Learning and Instruction*, *37*, 50-55.

Vamvakoussi, X., Van Dooren, W., & Verschaffel, L. (2012). Naturally biased? In search for reaction time evidence for a natural number bias in adults. *The Journal of Mathematical Behavior, 31*(3), 344-355.

Vamvakoussi, X., & Vosniadou, S. (2004). Understanding the structure of the set of rational numbers: A conceptual change approach. *Learning and Instruction, 14*(5), 453-467.

Vamvakoussi, X., & Vosniadou, S. (2010). How many decimals are there between two fractions? Aspects of secondary school students' understanding of rational numbers and their notation. *Cognition and Instruction,* 28(2), 181e209. http:// dx.doi.org/10.1080/07370001003676603.

Van Hoof, J., Degrande, T., Ceulemans, E., Verschaffel, L., & Van Dooren, W. (2018). Towards a mathematically more correct understanding of rational numbers: A longitudinal study with upper elementary school learners. *Learning and Individual Differences, 61*, 99-108.

Van Hoof, J., Janssen, R., Verschaffel, L., & Van Dooren, W. (2015). Inhibiting natural knowledge in fourth graders: towards a comprehensive test instrument. *ZDM*, *47*(5), 849-857.

Van Hoof, J., Lijnen, T., Verschaffel, L., & Van Dooren, W. (2013). Are secondary school students still hampered by the natural number bias? A reaction time study on fraction comparison tasks. *Research in Mathematics Education*, *15*(2), 154-164.

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*(1), 39-56.

VanLehn, K., Jones, R. M., & Chi, M. T. (1992). A model of the self-explanation effect. *The Journal of the Learning Sciences*, *2*(1), 1-59.

Vosniadou, S. (2014). Examining cognitive development from a conceptual change point of view: The framework theory approach. *European Journal of Developmental Psychology*, *11*(6), 645-661.

Vosniadou, S., & Verschaffel, L. (2004). Extending the conceptual change approach to mathematics learning and teaching. *Learning and Instruction, 14*(5), 445e548. http://dx.doi.org/10.1016/j.learninstruc.2004.06.014.

Wall, J., Thompson, C., & Morris, B. J. (2015). Confidence judgments and eye fixations reveal adults’ fractions knowledge. In Noelle, D. C., Dale, R., Warlaumont, A. S., Yoshimi, J., Matlock, T., Jennings, C. D., & Maglio, P. P. (Eds.), Proceedings of the 37th Annual Meeting of the Cognitive Science Society (pp. 2571–2576). Austin, TX: Cognitive Science Society.

Wilson, T. D. & Nisbett, R. E. (1978). The accuracy of verbal reports about the effects of stimuli on evaluations and behavior. *Social Psychology*, 41(2), 118-131.

Zhang, L., Fang, Q., Gabriel, F. C., & Szücs, D. (2014). The componential processing of fractions in adults and children: Effects of stimuli variability and contextual interference. *Frontiers in Psychology*, *5*, 981.

Appendix A

Stimuli used in Experiments 1 and 2.

Table A1.

Fraction stimuli in the number line estimation task from Experiment 1 (top) and 2 (Bottom).

Small component		Large Component	
SC1	SC2	LC1	LC2
1/5	3/15	5/25	6/30
2/9	4/18	6/27	8/36
1/4	3/12*	8/32	9/36
1/3	3/9*	10/30	12/36
3/7	9/21	12/28	15/35
1/2	2/4	12/24	15/30
5/9	10/18	15/27	20/36
2/3	4/6	6/9	8/12
3/4	6/8	9/12	12/16
4/5	8/10	12/15	20/25
5/6	10/12	15/18	20/24
Experiment 2			
1/6		12/72	
2/7[+]		24/84	
4/9		36/81	
3/5		27/45	
5/8		35/56	
6/7		54/63	

Note. Fractions are equivalent in magnitude in each row (e.g., 1/5 = 3/15 = 5/25 = 6/30). Stimuli with an asterisk are the two fractions that were changed from the original stimuli in Braithwaite and Siegler (2017).The original small component fraction 6/24 was changed to 3/12 and the original small component fraction 8/24 was changed to 3/9. [+]The fraction 2/7 was not used by Braithwaite and Siegler (2017).

Table A2. Fraction equivalence statements in Experiment 1.

True						False					
1/5	=	3/15	1/2	=	15/30	1/3	=	9/21	8/36	=	3/7
4/18	=	2/9	20/36	=	5/9	6/8	=	1/2	2/3	=	15/35
1/4	=	3/12	2/3	=	8/12	4/18	=	2/3	1/3	=	6/30
3/9	=	1/3	12/16	=	3/4	2/9	=	3/15	15/30	=	5/9
3/7	=	9/21	4/5	=	20/25	3/7	=	10/18	12/36	=	1/2
1/2	=	2/4	5/6	=	20/24	10/12	=	3/4	12/16	=	2/9
10/18	=	5/9	6/30	=	3/15	5/6	=	3/9	10/12	=	9/36
4/6	=	2/3	8/36	=	4/18	8/10	=	1/4	3/15	=	20/25
6/8	=	3/4	9/36	=	3/12	3/12	=	4/5	9/21	=	15/30
8/10	=	4/5	3/9	=	12/36	4/6	=	1/5	3/12	=	6/30
5/6	=	10/12	9/21	=	15/35	2/4	=	5/9	12/36	=	10/18
6/30	=	1/5	2/4	=	15/30	4/5	=	8/12	4/6	=	8/36
8/36	=	2/9	10/18	=	20/36	20/36	=	5/6	20/24	=	6/8
9/36	=	1/4	4/6	=	8/12	20/24	=	1/5	20/36	=	8/10
1/3	=	12/36	12/16	=	6/8	3/4	=	9/36	2/4	=	12/16
3/7	=	15/35	20/25	=	8/10	1/4	=	20/25	3/9	=	15/35
20/24	=	10/12				4/18	=	8/12			

Note. Statements were presented in a randomized order to participants.

Table A3.

Fraction comparison items by trial type. Bolded fractions are the fractions with the larger magnitudes and correct responses

Small Distance Comparisons			
Consistent Comparisons		Inconsistent Comparisons	
12/28	2/5	**4/7**	13/23
9/15	**28/41**	32/43	**9/11**
12/13	10/11	**11/14**	15/21
24/49	11/23	35/41	**11/12**
3/8	**11/27**	16/21	**7/9**

Large Distance Comparisons			
Consistent Comparisons		Inconsistent Comparisons	
2/5	**14/18**	12/29	**6/7**
34/43	12/33	**9/13**	14/45
22/45	2/21	11/23	**8/9**
11/19	**21/22**	17/46	**12/16**
13/17	3/8	**10/12**	13/29

Note. Small distance comparisons had a decimal distance between the fractions less than .14 (M = .04), and large distance comparisons had a distance greater than .37 (M = .41). Correct responses are bolded.

Table A4. Fraction equivalence statements from Experiment 2.

True			**False**		
12/72	=	1/6	1/6	=	36/81
2/7	=	24/84	54/63	=	2/7
4/9	=	36/81	4/9	=	35/56
27/45	=	3/5	3/5	=	24/84
35/56	=	5/8	5/8	=	27/45
6/7	=	54/63	12/72	=	6/7
3/18*	=	2/12*	16/36*	=	2/12*
12/27*	=	8/18*	6/21*	=	18/42*
12/14*	=	18/21*	6/30*	=	9/15*

Note. Fractions with an asterisk were not presented during the number line estimation task.

Statements were presented in a random order.

Appendix B

Full statistical models from Experiments 1 and 2.

Table B1.

Effects of component size and individual differences on raw estimates in Experiment 1.

Within subjects effects	*F*[a]	*MS*	*p*	η_p^2
Component size	0.52	0.001	0.604	0.007
Component Size * Equivalence	0.27	0.001	0.77	0.004
Component Size * Inhibition	0.27	0.001	0.776	0.004
Component size * Updating	2.18	0.004	0.115	0.03
Component Size * Condition	7.71	0.015	**0.001**	0.099
Component Size * Condition * Equivalence	8.04	0.016	**<.001**	0.103
Component Size * Condition * Inhibition	2.55	0.005	0.079	0.035
Component Size * Condition * Updating	0.95	0.002	0.394	0.013
Between subjects effects	*F*[b]	*MS*	*P*	η_p^2
Condition	< 0.01	< 0.001	0.968	< 0.001
Equivalence	0.01	< 0.001	0.979	< 0.001
Inhibition	0.43	0.005	0.512	0.006
Updating	1.63	0.019	0.206	0.023
Condition * Equivalence	4.51	0.052	0.037	0.061
Condition * Inhibition	0.33	0.004	0.565	0.005
Condition * Updating	12.55	0.144	< .001	0.152

Note. MS = Mean squared. [a]Within-subjects effects, Greenhouse-Geiser corrected degrees of freedom for F-statistic = 2.085, 145.918. [b]Between-subjects effects degrees of freedom for F-statistic = 1, 70.

Table B2.

Effects of component size and individual differences on PAE estimates in Experiment 1.

Within-subjects effects on PAE	F^a	*MS*	*p*	η_p^2
Component size	1.49	0.001	0.219	0.021
Component Size * Equivalence	1.87	0.001	0.135	0.026
Component Size * Inhibition	0.02	<.001	0.995	0
Component size * Updating	0.32	<.001	0.813	0.005
Component Size * Condition	3.58	0.003	0.015	0.049
Component Size * Condition * Equivalence	4.28	0.004	0.006	0.058
Component Size * Condition * Inhibition	0.35	<.001	0.789	0.005
Component Size * Condition * Updating	2.26	0.002	0.082	0.031

Between Subjects Effects	F^b	*MS*	*p*	η_p^2
Condition	1.60	0.026	0.211	0.022
Equivalence	19.29	0.316	< 0.001	0.216
Inhibition	1.76	0.029	0.189	0.025
Updating	5.79	0.095	0.019	0.076
Condition * Equivalence	1.79	0.029	0.186	0.025
Condition * Inhibition	0.48	0.008	0.491	0.007
Condition * Updating	0.01	< 0.001	0.897	< 0.001

Note. MS = mean squared. [a]Within-subjects effects degrees of freedom for F-statistic = 3, 210.

[b]Between-subjects effects degrees of freedom for F-statistic = 1, 70.

Table B3.

Effects of consistency and distance on fraction comparison accuracy in Experiment 1.

Within-subjects effects	F^a	*MS*	*p*	η_p^2
Consistency	6.857	0.653	0.011	0.083
Consistency * Condition	0.101	0.01	0.752	0.001
Distance	43.415	2.068	< 0.001	0.364
Distance * Condition	0.033	0.002	0.856	< 0.001
Consistency * Distance	21.953	0.498	< 0.001	0.224
Consistency * Distance * Condition	0.785	0.018	0.378	0.01

Between-subjects effects	F^a	*MS*	*p*	η_p^2
Condition	1.18	0.131	0.280	0.015

Note. MS = mean squared. [a]Within-subjects effects degrees of freedom for F-statistics= 1, 76.

Table B4.

Effects of component size and individual differences on PAE in Experiment 2.

Tests of Within-Subjects Effects	F^a	*MS*	*p*	η_p^2
Component size	6.29	0.007	0.014	0.061
Component size * Equivalence	5.84	0.007	0.018	0.057
Component size * Condition	0.20	< 0.001	0.654	0.002
Component size * Condition * Equivalence	0.212	< 0.001	0.646	0.002

Tests of Between-Subjects Effects	F^a	*MS*	*p*	η_p^2
Equivalence	44.41	0.396	<.001	0.316
Condition	0.01	<.001	0.946	< 0.001
Condition * Equivalence	0.05	<.001	0.829	< 0.001

Note. MS = mean squared. [a]Degrees of freedom for F-statistic = 1, 96.

Table B5.

Effects of component size and individual differences on raw estimates in Experiment 2.

Tests of Within-Subjects Effects	F^a	*MS*	*p*	η_p^2
Component size	1.82	0.005	0.18	0.019
Component size * Equivalence	1.58	0.005	0.213	0.016
Component size * Condition	0.26	0.001	0.612	0.003
Component size * Condition * Equivalence	0.34	0.001	0.561	0.004

Tests of Between Subjects effects	F^a	*MS*	*p*	η_p^2
Equivalence	2.21	0.012	0.141	0.022
Condition	9.79	0.051	0.002	0.093
Condition * Equivalence	8.60	0.045	0.004	0.082

Note. MS = mean squared. [a]Degrees of freedom for F-statistic = 1, 96.

Table B6.

Effects of component size and individual differences on PAE in Experiment 2 after filtering out participants who did not have codable strategy data.

Tests of Within-Subjects Effects	F^a	*MS*	*p*	η_p^2
Component size	5.68	0.005	0.02	0.07
Component size * Equivalence	5.24	0.005	0.025	0.065
Component size * Condition	0.01	<.001	0.946	< 0.001
Component size * Condition * Equivalence	< 0.01	<.001	0.946	< 0.001
Tests of Between-Subjects Effects				
Equivalence	17.35	0.064	<.001	0.188
Condition	0.02	<.001	0.882	< 0.001
Condition * Equivalence	< 0.01	<.001	0.958	< 0.001

Note. MS = mean squared. [a]Degrees of freedom for F-statistic = 1, 75.

Table B7.

Effects of component size and individual differences on raw estimates in Experiment 2 after filtering out participants who did not have codable strategy data.

Tests of Within-Subjects Effects	F^a	*MS*	*p*	η_p^2
Component size	0.568	0.001	0.453	0.008
Component size * Equivalence	0.461	0.001	0.499	0.006
Component size * Condition	4.643	0.009	0.034	0.058
Component size * Condition * Equivalence	3.566	0.007	0.063	0.045
Tests of Between-Subjects Effects				
Equivalence	0.059	0.000	0.809	0.001
Condition	4.03	0.009	0.048	0.051
Condition * Equivalence	3.73	0.008	0.057	0.047

Note. MS = mean squared. [a]Degrees of freedom for F-statistic = 1, 75.