

# Comparing Fraction Magnitudes: Adults' Verbal Reports Reveal Strategy Flexibility and Adaptivity, but Also Bias

Andreas Obersteiner<sup>1,2</sup>, Martha Wagner Alibali<sup>2</sup>, Vijay Marupudi<sup>3</sup>

[1] TUM School of Social Sciences and Technology, Technical University of Munich, Munich, Germany. [2] Department of Psychology, University of Wisconsin-Madison, Madison, WI, USA. [3] Georgia Institute of Technology, Atlanta, GA, USA.

Journal of Numerical Cognition, 2022, Vol. 8(3), 398–413, <https://doi.org/10.5964/jnc.7577>

Received: 2021-09-29 • Accepted: 2022-08-01 • Published (VoR): 2022-11-16

Handling Editor: Jake McMullen, University of Turku, Turku, Finland

Corresponding Author: Andreas Obersteiner, Heinz Nixdorf Chair of Mathematics Education, TUM School of Social Sciences and Technology, Technical University of Munich, Arcisstr. 21, D-80333 Munich. E-mail: andreas.obersteiner@tum.de

Related: This article is part of the JNC Special Issue “Mathematical Flexibility”, Guest Editors: Marian Hickendorff, Jake McMullen, & Lieven Verschaffel, Journal of Numerical Cognition, 8(3), <https://doi.org/10.5964/jnc.v8i3>

Supplementary Materials: Materials [see [Index of Supplementary Materials](#)]



## Abstract

Many studies have used fraction magnitude comparison tasks to assess people's abilities to quickly assess fraction magnitudes. However, since there are multiple ways to compare fractions, it is not clear whether people actually reason about the holistic magnitudes of the fractions in this task and whether they use multiple strategies in a flexible and adaptive way. We asked 72 adults to solve challenging fraction comparisons (e.g.,  $31/71$  vs.  $13/23$ ) on a computer. In some of these comparisons, using benchmarks (i.e., reference numbers such as  $1/2$ ) was potentially beneficial. After each trial, participants provided verbal reports of their strategies. We found that participants used a large variety of strategies. The majority of strategies were holistic and relied on fraction magnitudes, and most of these strategies were based on benchmarks. Participants sometimes used gap comparison (i.e., comparing the differences between each fraction's numerator and denominator), a heuristic that is not always valid and that does not rely on fraction magnitudes. Participants used strategies flexibly: they used many different strategies, they used highly efficient strategies most often, and they adapted their strategy use to features of the items. However, participants sometimes used gap comparison on items for which it did not yield the correct response, and this lack of adaptivity partly explained the “natural number bias” observed in this study.

## Keywords

fraction comparison, strategy use, natural number bias, trial-by-trial assessment

How do people decide which of two fractions is larger? Although younger children often fail to compare even simple fractions accurately (Rinne et al., 2017), educated adults are usually quick and accurate, even when given challenging, unfamiliar fraction pairs (e.g.,  $31/71$  vs.  $13/23$ ; Barraza et al., 2017; DeWolf & Vosniadou, 2015; Morales et al., 2020). However, it is unclear what strategies adults use to achieve their high performance in mental fraction comparison. Specifically, it is not known to what extent people rely on overall fraction magnitudes or on other strategies that do not involve fraction magnitudes. There is substantial variability in strategy use within and between individuals in simple fraction comparisons (e.g., Reinhold et al., 2020; Rinne et al., 2017), but empirical evidence on more challenging items



with unfamiliar fractions is scarce. The aim of this study was, therefore, to assess individual strategy use on challenging fraction comparisons using a trial-by-trial approach.

Understanding strategy use in fraction comparison is important for many reasons. First, flexible and adaptive strategy use is an important facet of mathematical competency (Heinze et al., 2009; Siegler & Lemaire, 1997), and competence with fractions is associated with later mathematical achievement (Siegler et al., 2012). Thus, studying strategy use in fraction comparisons can contribute to better understanding of an important mathematical domain. Second, although there are several mathematically valid strategies for comparing fractions (i.e., clearly defined algorithms that yield a correct answer for any comparison), these strategies are often cognitively demanding, for example, when they involve two-digit multiplication. Consequently, people sometimes rely on heuristics that are not always valid or on estimations that are less precise. Heuristics can be correct and efficient in some cases, but they lead to incorrect responses in other cases, resulting in biased reasoning patterns (e.g., "natural number bias"; Ni & Zhou, 2005). Studying fraction comparison strategies with challenging fraction pairs therefore contributes to understanding people's (imperfect) decision making processes in mathematical problem situations (Fischbein, 1987; Gillard et al., 2009). Third, although there is evidence that people can construct fraction magnitude representations quickly, particularly for familiar fractions (Liu, 2018; Schneider & Siegler, 2010), little is known about how people do so for unfamiliar fractions. Finally, fraction comparison tasks are often used to assess fraction magnitude representations (e.g., Schneider & Siegler, 2010). However, the validity of this measure hinges on the assumption that people actually engage in magnitude processing when comparing fractions. This assumption has been found to be untenable for simple fraction comparisons (Obersteiner et al., 2013; Reinhold et al., 2020), but it is still under debate regarding challenging fraction comparisons (e.g., Obersteiner et al., 2020; Schneider & Siegler, 2010).

## Fraction Magnitude Processing

Studies of educated adults reveal that most can make fraction comparisons quickly and accurately (e.g., Barraza et al., 2017; DeWolf & Vosniadou, 2015; Meert et al., 2010; Morales et al., 2020; Obersteiner et al., 2013; Van Hoof et al., 2013). Across studies, participants' average accuracy was fairly high (around 90%) and response times were short (a few seconds), and this held true, not only for pairs with familiar fractions (e.g.,  $1/2$ ,  $3/4$ ) or common numerators (e.g.,  $10/27$  vs.  $10/21$ ) or denominators (e.g.,  $10/27$  vs.  $21/27$ ) (Meert et al., 2010; Van Hoof et al., 2013), but also for comparisons that were more challenging (e.g.,  $31/71$  vs.  $13/23$ ) (Barraza et al., 2017; DeWolf & Vosniadou, 2015; Morales et al., 2020; Obersteiner et al., 2013). Challenging items included less familiar fractions, fractions made up of double-digit integers, and fraction pairs with small numerical differences.

Studies of educated adults also often report pronounced "distance effects" in fraction comparisons—that is, strong associations between the numerical distance between the two fractions and participants' average accuracy and/or response times. Accuracy tends to be higher, and response times lower, with increasing numerical distance. Such distance effects suggest the processing of holistic (overall) magnitude representations, similar to those recruited for whole number magnitude comparisons (Moyer & Landauer, 1967). Obersteiner et al. (2013) found that numerical distance accounted for a substantial amount of the variance in response times for a set of challenging fraction comparisons (up to 71%, depending on the specific item subset), suggesting that people activate holistic magnitudes when comparing fractions. Notably, the distance effect appears to be more pronounced for challenging fraction comparisons than for simple ones, for which alternative short-cut strategies (e.g., comparing numerators) are applicable (Bonato et al., 2007; Ganor-Stern et al., 2011; Obersteiner et al., 2013; Reinhold et al., 2020).

One limitation of previous research is that almost all studies aggregated data across participants (e.g., Meert et al., 2010; Obersteiner et al., 2013; Van Hoof et al., 2013). Such analyses do not capture the potentially large variability in individual patterns of accuracy and response times. Substantial variability has frequently been documented in children, whose responses are often incorrect (Reinhold et al., 2020; Rinne et al., 2017). Although adults generally display less variability in aggregated performance than children, they may still vary in the specific *strategies* they use to solve challenging fraction comparisons. For familiar fractions, such as  $1/2$ , most adults may readily activate holistic magnitude representations (Liu, 2018); however, it is not well understood how people evaluate the magnitudes of less familiar fractions.

## Strategy Use in Fraction Comparison

Various facets of strategy use have been identified and characterized in previous work (Heinze et al., 2009; Lemaire & Siegler, 1995). *Strategy repertoire* refers to the set of different strategies that individuals know and can apply. In the case of fraction comparison, individuals may have several generally valid strategies for comparing fractions in their repertoires. For example, one can convert both fractions to the same denominator and then compare numerators ( $2/5 = 16/40$ ;  $5/8 = 25/40$ , hence  $5/8 > 2/5$ ); one can convert both fractions into decimals and compare them ( $2/5 = 0.4$ ;  $5/8 = 0.625$ , and  $0.625 > 0.4$ , so  $5/8 > 2/5$ ); or one can use  $1/2$  as a benchmark (i.e.,  $2/5$  is less than half,  $5/8$  is more than half, hence  $5/8 > 2/5$ ). An individual may know and use all three of these strategies or only a subset of them.

A second facet of strategy use is *strategy frequency*, which refers to how often people use particular strategies. Strategy frequency may depend on characteristics of the context or task. For example, people may use certain strategies more frequently when comparing fractions mentally than when using paper and pencil. Contextual factors (e.g., time pressure) may also encourage people to use strategies or heuristics that are not generally valid but that lead to correct responses on some comparisons. For example, the *gap comparison* strategy consists of finding the numerical *difference* (gap) between the numerator and denominator of each fraction and choosing the fraction with the *smaller* gap as the larger fraction. In some cases, this strategy yields the correct response (e.g.,  $31/71$  vs.  $13/23$ , where the larger fraction has a smaller gap [40 vs. 10]), but in other cases, it does not (e.g.,  $31/65$  vs.  $13/31$ , where the larger fraction has a larger gap [34 vs. 18]). Moreover, instructing or encouraging people to use specific strategies may also affect the frequency with which they use particular strategies (Fazio et al., 2016).

A third facet of strategy use is *strategy efficiency*, which refers to the speed and accuracy with which strategies are applied. For example, in challenging fraction comparisons, it may be more efficient to use benchmarks (i.e., reference numbers such as  $1/2$ ) than to convert fractions into decimals, because benchmarking relies on estimation instead of exact mental computation. In some situations, people may rely on generally invalid heuristics that are efficient because they are easy to apply (e.g., gap comparison or simple comparison of numerators), even though such heuristics carry the risk of yielding incorrect answers (see above).

Two additional, related facets of strategy use are *strategy flexibility* and *adaptivity*. *Flexibility* refers to the ability to shift between different strategies, and *adaptivity* refers to how frequently people choose strategies that are best suited for a given problem. Note that identifying which strategy is “best suited” is difficult because it requires both a normative perspective (i.e., which strategies are valid and applicable on a given problem) and an individual perspective (i.e., which strategies a given individual can apply quickly and accurately).

Assessing all facets of strategy use simultaneously is a methodological challenge. For example, assessing strategy efficiency requires recording accuracy and response times, but these measures alone do not allow for straightforward conclusions about specific strategies. Accordingly, previous studies have investigated fraction comparison strategies using a variety of methods, including identifying error patterns across sets of items (González-Forte et al., 2020; Stafylidou & Vosniadou, 2004), measuring reaction times (Gómez & Dartnell, 2019; Morales et al., 2020; Obersteiner et al., 2020), and studying eye movement patterns (Obersteiner & Tumpek, 2016). In all of these studies, data were aggregated over many participants and items, and the focus of the analyses was typically on the broad distinction between holistic and componential strategies. *Holistic* strategies are strategies that rely on reasoning about the magnitudes of fractions as wholes, whereas *componential* strategies rely on reasoning about fraction components (i.e., numerators and denominators), without considering holistic magnitudes. These past studies have shown that people are more likely to use componential strategies in simple comparisons, in which fractions have the same numerators or denominators (e.g.,  $13/25$  vs.  $21/25$ ), and use holistic strategies when this is not the case. However, studies using these methods did not characterize individual strategy use in detail.

One approach that allows for individual-level analyses of strategy use is collecting verbal reports on a trial-by-trial basis. However, previous studies using verbal reports have not considered challenging fraction comparisons. For example, Faulkenberry and Pierce (2011) collected trial-by-trial strategy reports from adults; however, the items included common fractions such as  $1/2$  and  $1/3$ , potentially affecting the range of strategies used. Clarke and Roche (2009) and González-Forte and colleagues (2019) collected trial-by-trial strategy reports from children in grades 6 and 7, respectively; however, both studies used item sets with very simple fraction comparisons. These studies provided some

empirical evidence that people use a range of strategies to make simple fraction comparisons, but they do not provide information about how people make difficult comparisons quickly.

One study that used a trial-by-trial analysis did include a few difficult comparisons. Fazio, DeWolf, and Siegler (2016) studied fraction comparison with verbal reports in university students who were presented with eight different types of fraction pairs. One type of comparison that they included was very challenging—pairs with two-digit, non-equal components and small numerical distances. On these difficult comparisons, participants often used *benchmark* strategies (referred to as “general magnitude reference”), *gap comparisons*, or strategies that were coded as “other” (infrequent or uncodable strategies). However, the experiment included only three items of this type. Thus, a larger set of items and a more fine-grained analysis of strategy use is needed to provide a more comprehensive picture of strategy use on challenging fraction comparisons.

## Strategy Use and the Natural Number Bias

Analyzing strategy use in fraction comparison may contribute to better understanding of the *natural number bias* that has been well documented in the literature (e.g., Alibali & Sidney, 2015; Ni & Zhou, 2005). One manifestation of the natural number bias is a systematic difference in performance (in accuracy and/or response times) between fraction comparisons that are *congruent* with naïve natural number reasoning (i.e., the larger fraction has the larger whole number components, e.g.,  $1/4$  vs.  $3/5$ ) and *incongruent* with naïve natural number reasoning (i.e., the larger fraction has the smaller whole number components, e.g.,  $1/4$  vs.  $2/9$ ). Research has documented pronounced individual differences in bias patterns (Gómez & Dartnell, 2019; González-Forte et al., 2020; Rinne et al., 2017). Although studies using *simple* fraction comparisons have often reported better performance on congruent than incongruent items (e.g., Van Hoof et al., 2013), studies that addressed individual differences or that included more challenging fraction pairs have frequently documented the reverse pattern (i.e., better performance on incongruent than congruent comparisons) (Barraza et al., 2017; DeWolf & Vosniadou, 2015; Obersteiner et al., 2020).

The performance difference between congruent and incongruent items has been attributed to people’s use of simple heuristics (e.g., “larger numerator/denominator makes larger fraction” or “smaller numerator/denominator makes larger fraction” heuristics; González-Forte et al., 2020; Reinhold et al., 2020). However, the sources of the reverse bias pattern that is sometimes observed in educated adults are less well understood. It seems unlikely that educated adults use a “smaller numerator/denominator makes larger fraction” heuristic, and some authors have suggested that other heuristics, especially gap comparison, could explain the reverse bias pattern (e.g., Gómez & Dartnell, 2019; Obersteiner et al., 2020). The reason is that for comparisons of fractions smaller than 1 (the item type used in most previous studies), gap comparison yields the correct response for all *incongruent* items, but it may or may not yield the correct response for *congruent* items. For example,  $31/71$  vs.  $13/23$  is an incongruent comparison, and applying gap comparison (40 vs. 10, hence  $13/23$  must be larger) yields the correct response. However,  $39/51$  vs.  $18/31$  and  $35/51$  vs.  $18/31$  are both congruent comparisons, and gap comparison leads to the correct response in the former case (gaps of 12 vs. 13) but not the latter (16 vs. 13). If people rely strongly on gap comparison and the item set includes both congruent and incongruent items, people would be more efficient (faster and more accurate) on incongruent than on congruent items, which might partly explain the reverse bias pattern.

## The Present Study

The aim of the present study was to assess adults’ strategy use on challenging fraction comparisons. We considered the facets of strategy use described earlier. Considering participants’ *strategy repertoires* and *frequencies*, we focus in particular on the use of *holistic* strategies (especially the use of benchmarks) and *componential* strategies (especially the use of gap comparison). We expected that participants would use holistic strategies to a large extent, because previous studies have documented adults’ strong performance and distance effects in fraction comparison (see section entitled Strategy Use in Fraction Comparison), both of which are in line with holistic strategy use. We were less certain about whether adults would use componential strategies, especially gap comparison.

Regarding *strategy efficiency*, we did not make specific predictions. Although holistic strategies may be faster to apply than some componential strategies, and they may consistently yield accurate answers, they also include approximations that may be imprecise in challenging fraction comparisons.

Regarding *strategy flexibility and adaptivity*, we expected that individual participants would use many different strategies, as in previous research with simple comparisons (e.g., Fazio et al., 2016). We also expected that strategy use would depend on features of the items. Specifically, we expected that participants would use benchmark strategies more frequently when a benchmark was in between the given fractions (e.g., when one fraction was larger and the other smaller than  $1/2$ ), or when one fraction was particularly close to a benchmark.

Finally, we were interested in whether we could evoke shifts in participants' strategy use by highlighting the usefulness of a holistic strategy. To this end, we encouraged some participants to use benchmark strategies. If people can easily shift towards more frequent use of holistic strategies, it would suggest that strategy use is malleable, which would have implications for educational practice.

## Method

### Participants

Participants were 72 undergraduates (32 male, 37 female, 1 non-binary, 2 not reported;  $M = 19.30$  years,  $SD = 1.32$ ; 50% White, 39% Asian, 4% Black, 1% Hispanic/Latinx, 3% biracial, 1% multiracial, 1% other) from a moderately selective university in the Midwestern United States. Participants received extra credit in Introductory Psychology or compensation of \$10 for their participation. Participants were randomly assigned either to receive ( $n = 40$ ) or not receive ( $n = 32$ ) a tip that using benchmarks could be helpful in comparing fractions (see Procedure).

### Stimuli

We used the same set of 56 fraction comparison items as in a previous study that did not involve strategy reports (Obersteiner et al., 2020). All fractions were smaller than 1 and in simplest form. All denominators and most numerators (86%) were two-digit numbers. The two numerators or two denominators within a fraction pair were never equal, and they were never integer multiples of one another. Half of the items were congruent (i.e., the larger fraction had a larger numerator and larger denominator than the smaller fraction) and half were incongruent (i.e., the larger fraction had a smaller numerator and smaller denominator than the smaller fraction). Within the congruent and the incongruent subsets, there were three types of items with respect to the benchmarks 0,  $1/4$ ,  $1/2$ ,  $3/4$ , and 1. The first type was *straddle* items, in which one fraction was smaller and the other larger than one of the benchmarks, so that fraction pairs of this type "straddled" either  $1/4$ ,  $1/2$ , or  $3/4$ . The second type was *in-between* items, in which both fractions' magnitudes were between two adjacent benchmarks but did not straddle a benchmark. Thus, the two fractions were either both larger than  $1/4$  and smaller than  $1/2$ , or both larger than  $1/2$  and smaller than  $3/4$ . The third type was *close-to-0-or-1* items, in which both fractions were either smaller than  $1/4$  (and larger than 0), or larger than  $3/4$  (and smaller than 1). Even though close-to-0-or-1 items were also between benchmarks, we considered them a separate category because previous research suggested that fractions' proximity to 0 or 1 made these items easier to solve (Obersteiner et al., 2020). All items had small numerical distances ( $M = 0.14$ ; range 0.09–0.16), and all fractions had similar distances to the closest benchmark (0,  $1/4$ ,  $1/2$ ,  $3/4$ , or 1;  $M = 0.06$ ; range 0.03–0.11).

The item set was split into two subsets of 28 items each, such that each subset contained equally many items of each type, as described above. We added three filler items to each subset, and these fillers were not included in the analysis. Filler items were easy fraction pairs (with same numerators or with very large differences between fractions) that did not fall into any of the categories described above. These items were added to maintain participants' motivation and attention.<sup>1</sup> Thus, each participant was presented with a total of 31 items, consisting of one subset of 28 items and three

---

1) Performance on items that followed filler items was comparable to performance on other items in speed, accuracy, and strategy use.



filler items. Preliminary analyses revealed no differences in performance across the two subsets, so they were collapsed for analysis.

## Procedure

Data were collected in small group sessions, with each participant working individually at a computer. Items were presented using E-Prime software. Participants wore headphones with microphones. They were presented with two fractions at a time and asked to choose the greater fraction as quickly and accurately as possible by pressing the corresponding left (“f”) or right (“j”) key within 15 seconds. After participants pressed the key, the fraction pair remained visible on the screen, and the question “How did you solve this problem?” appeared above the fraction pair. Participants then had another 20 seconds to respond by speaking into their microphones. After the given time limit, the next item appeared automatically.

There were two practice items before the experiment started. Accuracy feedback was provided for practice items but not for test items. Test items were presented in random order, and the position of the larger fraction (left or right) was counterbalanced across trials.

Participants in the tip condition saw an additional screen after the general instructions. The text on the screen suggested that it could be helpful to think of numbers such as  $1/2$ ,  $1/4$ , or  $3/4$  that could be used as benchmarks to compare fractions. The example “ $5/8$  vs.  $3/7$ ” was provided to illustrate the benchmark strategy. The text said that one could think that  $5/8$  is larger than  $1/2$ , and that  $3/7$  is smaller than  $1/2$ , so that  $5/8$  must be the larger fraction. No further explanation was provided.

## Data Analysis

### Accuracy and Response Time Data

We first provide an overview of accuracy and response time data (previously reported in Obersteiner et al., 2019). The data were analyzed in SPSS 23 using a General Estimation of Equations (GEE) model that accounts for repeated measures within subjects. In these analyses, the within-subject factors were *congruency* (congruent/incongruent) and *item type* (close-to-0-or-1/straddle/in-between); the between-subject factor was *tip* (yes/no). For the dichotomous accuracy data, we used a binary logistic regression model. For response time data, we used a linear regression model with a logarithmic link function.

### Verbal Responses

Participants’ verbal responses ( $N = 2016$ ) were transcribed and then coded by two coders (the first and third authors of this paper). The coding scheme was developed following a combined deductive and inductive approach: We first defined the major categories (holistic and componential) and then identified specific strategies that fell into these categories. Specific strategies (e.g., benchmark, gap comparison) were identified based on the literature (e.g., Clarke & Roche, 2009; Fazio et al., 2016). We modified and extended the coding scheme as necessary until no new strategies were observed in the data. Table 1 provides an overview of strategies that were used on at least 1% of items.

Verbal recordings from 18 randomly-selected participants ( $n = 504$  responses, 25% of the data) were double coded, and interrater agreement was evaluated for this subset. Agreement was high, both for coding strategies into broad categories (holistic vs. componential), Cohen’s kappa = 0.84, and for coding strategies into subcategories, kappa = 0.77. Disagreements were discussed until agreement was reached, and the agreed-upon codes were used in the analysis.

Participants sometimes reported two strategies on the same trial. In most cases, one strategy was clearly the primary strategy, and another was mentioned only briefly (e.g., as a possible approach). In these cases, we coded only the primary strategy. Otherwise (in 17% of all trials), we coded both strategies.<sup>2</sup>

Because our primary interest was the distinction between holistic and componential strategies, we considered as holistic all strategies in which people relied on overall fraction magnitudes. Importantly, categorizing a strategy as

---

2) Note that for these reasons the total number of strategies in the following analyses is larger than the number of verbal responses coded.

holistic did not imply that people needed to *directly* activate a holistic magnitude representation; they could also *construct* a holistic magnitude representation.

**Table 1***Fraction Comparison Strategies*

Category / Strategy	Description	Sample Response
<b>Holistic</b>		
Benchmark	Compare one or both fractions to another number (e.g., to $\frac{1}{2}$ )	[11/26 vs. 34/59] The right fraction is a little bit more than half while the left is less than half.
Multiplicative reasoning	Reason about the multiplicative relationship between numerator and denominator of each fraction	[6/19 vs. 19/96] 6 goes into 19 a little more than 3 times whereas 19 goes into 96 probably 6 or so times, maybe 5, so 6 over 19 is bigger.
Division/Conversion	Perform a division and/or convert one or both fractions to a decimal or percentage	[22/31 vs. 37/64] If you divide 22 by 31 it is a larger number than when you divide 37 by 64, by a great amount.
<b>Componential</b>		
Gap comparison*	Compare the differences (gaps) between numerator and denominator of each fraction	[68/83 vs. 49/52] 49 is closer to 52 than 68 is to 83.
Multiplication	Multiply one fraction to get similar numerators or denominators	[17/31 vs. 11/16] 11 over 16 can be 22 over 32 and 22 is much greater than 17, so the second one is greater.
Component comparison*	Compare fraction components without calculation	[22/49 vs. 13/41] 22 over 49 because the numerator is bigger.
Component approximation	Note that the numerators or the denominators are nearly equal, so compare the other two components	[17/91 vs. 13/41] 17 and 13 are close but 91 is way bigger than 41, so 13 over 41 is clearly bigger than 17 over 91.
Multiplicative comparison	Compare the multiplicative relation between the two numerators with that between the two denominators	[32/57 vs. 11/27] The top number is like tripled while the bottom number only doubled so the one on the left is gonna be bigger.
<b>Other</b>		
Unintelligible or unclear	Response is unintelligible or unclear	-
Missing	No response or response unrelated to task	-
Statement/Knowing	State the response without reasoning or by simply knowing	[13/43 vs. 6/37] 13 out of 43 is larger than 6 out of 37.
Guessing/Intuition	Guess or refer to intuition	Just guessing.
Don't know	State that one does not know the answer	I don't know.

Note. Heuristics that are not generally valid are marked with an asterisk (\*).

## Results

### Accuracy and Response Times

Average accuracy (Acc;  $M = 86\%$ ,  $SD = 11\%$ ) and response times (RT;  $M = 4813$  ms,  $SD = 1432$ ) were comparable with previous studies, suggesting that participants were well able to solve these challenging fraction comparisons. Table 2 displays the GEE model statistics for accuracy and response times. Participants were significantly more accurate and faster on incongruent than congruent items, indicating a *reverse bias* pattern (Acc:  $M = 94\%$ ,  $SE = 1\%$  vs.  $M = 81\%$ ,  $SE = 2\%$ ;  $OR = 1.14$ , 95% CI [1.09, 1.20]; RT:  $M = 4285$ ,  $SE = 209$  vs.  $M = 4652$ ,  $SE = 175$  ms; mean difference = 367 ms, 95% CI [149, 586]).

**Table 2**

Results of the GEE Procedure on Accuracy and Response Times

Source	Accuracy			Response Time		
	Wald $\chi^2$	df	p	Wald $\chi^2$	df	p
Congruency	38.86	1	< .001	11.59	1	< .001
Item Type	24.63	2	< .001	74.63	2	< .001
Tip	0.15	1	.695	0.00	1	.969
Congruency x Item Type	2.89	2	.236	0.23	2	.892
Item Type x Tip	5.46	2	.065	1.76	2	.416
Congruency x Tip	0.29	1	.592	4.51	1	.034
Congruency x Item Type x Tip	1.03	2	.598	0.80	2	.669

There was a main effect of *item type* for both accuracy and response times. Participants were most accurate and fastest on close-to-0-or-1 items ( $M = 93\%$ ,  $SE = 1\%$ ;  $M = 3917$  ms,  $SE = 178$ ), followed by straddle items ( $M = 88\%$ ,  $SE = 2\%$ ;  $M = 4744$  ms,  $SE = 217$ ) and then in-between items ( $M = 83\%$ ,  $SE = 2\%$ ;  $M = 4789$  ms,  $SE = 192$ ); planned contrasts revealed that all pairwise comparisons were significant ( $ps < .05$ ; OR 1.05–1.12; RT differences 113–121) except for the difference in response times between the latter two item types ( $p = .705$ , RT difference = 118, 95% CI [-276, 187]). There was no main effect of *tip*, either for accuracy or response times. For response times, but not accuracy, the effect of congruency was qualified by an interaction with *tip*. Participants were, overall, faster on incongruent than on congruent items, but this difference was smaller (and nonsignificant) for participants who received the tip ( $M = 4403$ ,  $SE = 229$ , vs.  $M = 4542$ ,  $SE = 280$  ms,  $p = .349$ , mean diff = 138, 95% CI [-151, 428]) than for participants who did not ( $M = 4170$ ,  $SE = 263$ , vs.  $M = 4766$ ,  $SE = 311$  ms,  $p < .001$ ; mean diff = 596, 95% CI [270, 922]). Thus, the tip somewhat reduced the effect of item congruency for response times, but not for accuracy. No other interactions were significant.

## Strategy Use

One participant was excluded from the strategy analyses due to technical problems with the recording. As reported above, the effect of the tip about benchmark strategies on accuracy and response times was generally small, and the main effect of tip was not significant for either accuracy or response times. This was also the case for strategy use. Participants who received the tip used holistic strategies only slightly more often (52%) than participants who did not receive the tip (49%), and this difference was not significant,  $\chi^2(1) = 2.46$ ,  $p = .117$ . Therefore, for the remainder of the analyses, we collapsed the tip groups.<sup>3</sup>

## Strategy Repertoire and Frequency

Participants used many different strategies (see Table 1). As seen in Table 3, overall, participants used holistic strategies more frequently (51% of responses) than componential strategies (38%).

The large majority of holistic strategies (82%) involved benchmarks. Benchmarks were either used as numbers straddled by the given fractions (“one fraction larger, the other smaller”; 19%) or, more often, as reference numbers to which one or both fractions were compared (“fraction close to benchmark”; 81%). The most prominent benchmarks were 1/2 (used in 42% of benchmark strategies), 1 (29%), 1/3 (10%) and 1/4 (4%). Several other numbers (including 0, 3/4 and 2/3) were infrequently used as benchmarks, as well. Holistic strategies that did not include benchmarks involved either multiplicative reasoning about the ratio between numerator and denominator or performing a division, often to convert fractions to an exact decimal or percentage. These strategies were relatively infrequent.

3) We ran all analyses reported in the following separately for participants in the tip and no-tip group. All differences between the two groups were small and did not change the overall conclusions. For separate analyses, see Supplementary Materials.



**Table 3***Percent Use of Each Fraction Comparison Strategy*

Category / Strategy	Percent Use Within Category	Percent Use Overall
<b>Holistic</b>		<b>51</b>
Benchmark	82	42
Multiplicative reasoning	15	7
Division/Conversion	3	2
<b>Componential</b>		<b>38</b>
Gap comparison	43	16
Multiplication	23	9
Component comparison	14	5
Component approximation	11	4
Multiplicative comparison	4	2
Other	5	2
<b>Other</b>		<b>12</b>
Unintelligible or unclear	40	5
Missing	30	4
Statement/Knowing	13	2
Guessing/Intuition	10	1
Don't know	6	1

Among componential strategies, the most frequent was gap comparison, which was the second-most-frequent strategy overall. Other componential comparisons were multiplication strategies, in which fraction components were multiplied to get common (or similar) numerators or denominators. It is notable that component comparison (i.e., simple comparison of numerators or denominators) was used infrequently (5% overall).

### Strategy Efficiency

Although mean accuracy was high overall, holistic strategies led to correct responses slightly more often ( $M = 90\%$ ,  $SE = 30$ ) than componential strategies ( $M = 86\%$ ,  $SE = 35$ ), Wald  $\chi^2(1) = 4.47$ ,  $p = .035$ ,  $OR = 1.04$ , 95% CI [1.01, 1.08]. There was no significant difference in response times between holistic and componential strategies,  $\chi^2(1) = 1.92$ ,  $p = .166$ . Participants who used holistic strategies more frequently tended to be more accurate overall; percentage use of holistic strategies was correlated with accuracy,  $r(69) = 0.29$ ,  $p = .014$ , but not with average response times,  $r(69) = -0.01$ ,  $p = .976$ .

To evaluate how efficiently (i.e., quickly and accurately) different strategies were used, we calculated the Inverse Efficiency Score (IES), which combines accuracy and response times into one measure (Bruyer & Brysbaert, 2011). The IES is calculated by dividing average response times by accuracy, and it can be interpreted as the average response time needed to provide a correct response (i.e., smaller values indicate higher efficiency). Table 4 displays accuracy, response times, and IES for each strategy. Among holistic strategies, benchmarking was the most efficient. Among componential strategies, component approximation was most efficient, followed by gap comparison. Comparing benchmark and gap comparison (the two most frequent strategies), gap comparison was more efficient.

### Strategy Flexibility and Adaptivity

Across the 28 comparisons, participants used, on average, five of the strategies described in Table 3 ( $M = 5.13$ ,  $SD = 1.22$ ; range: 2–8).<sup>4</sup> No participant rigidly used a single strategy. Thus, participants used strategies flexibly. The number

4) Note that this analysis did not include strategies that were coded as “other”.

of different strategies participants used was not correlated with their overall accuracy,  $r(69) = 0.11$ ,  $p = .373$ , or response times,  $r(69) = 0.00$ ,  $p = .988$ .

Participants used strategies adaptively in the sense that they frequently used highly efficient strategies—primarily benchmarking and gap comparison (ranked fourth and second, respectively, in Table 4). It is notable that participants rarely used the most efficient strategy, component approximation, presumably because it was applicable only on a subset of the items (i.e., those with similar numerators or denominators).

**Table 4**

*Accuracy, Response Times, and Efficiency Scores for all Strategies*

Category / Strategy	Accuracy (%)	Response Time (ms)	IES (ms)
<b>Holistic</b>	<b>90</b>	<b>4812</b>	<b>53.5</b>
Benchmark	90	4695	52.2
Multiplicative reasoning	90	5374	59.7
Division/Conversion	87	5409	62.2
<b>Componential</b>	<b>86</b>	<b>4718</b>	<b>54.9</b>
Gap comparison	85	4122	48.5
Multiplication	87	6155	70.7
Component comparison	77	4845	62.9
Component approximation	98	4072	41.6
Multiplicative comparison	95	4870	51.3

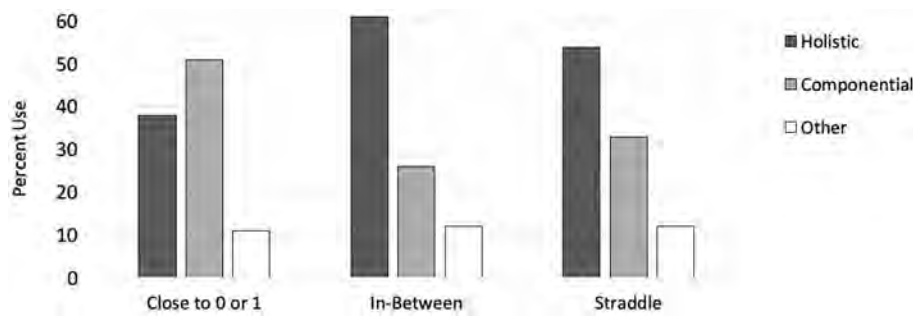
*Note.* IES (Inverse Efficiency Score) = Response Time / Accuracy, with lower values representing higher efficiency.

We also evaluated whether strategy use varied depending on item type. First, we considered the three item types (straddle, in-between, and close-to-0-or-1). Recall that participants were most accurate and fastest (i.e., most efficient) on close-to-0-or-1 items, followed by straddle items, and then in-between items (see section entitled Accuracy and Response Times). A major difference between the three item types was that participants used holistic strategies much more frequently than componential strategies for in-between, 60% vs. 27%;  $\chi^2(1) = 82.51$ ,  $p < .001^5$ , and straddle items, 53% vs. 34%;  $\chi^2(1) = 44.15$ ,  $p < .001$ , but the opposite was true for close-to-0-or-1 items, 38% vs. 54%;  $\chi^2(1) = 16.70$ ,  $p < .001$ . Figure 1 illustrates this difference, and Figure 2 presents data for all strategies for each of the three item types. On in-between and straddle items, participants frequently used benchmarking, and they did so more often than they did on close-to-0-or-1 items (Wald  $\chi^2(2) = 47.85$ ,  $p < .001$ ; all direct comparisons between the three types  $p < .001$ ,  $OR = 1.07$ – $1.19$ ). On close-to-0-or-1 items, participants still used benchmarking most often. However, on close-to-0-or-1 items, participants also used gap comparison relatively frequently, and they did so more often than on the other two item types (Wald  $\chi^2(2) = 50.31$ ,  $p < .001$ ; all direct group comparisons  $p < .05$ ,  $OR = 1.03$ – $1.13$ ). Participants also used multiplicative reasoning less frequently on close-to-0-or-1 items than on the other item types (4% vs. 9% vs. 9%) and they used component approximation (the most efficient strategy, see Table 4), more frequently (10% vs. 1% vs. 3%), although these strategies were used infrequently overall. In sum, participants adapted their strategy use to item type, with benchmarking playing a special role on in-between and straddle items.

5) Non-parametric chi-square test of the distributions of holistic vs. componential strategies.

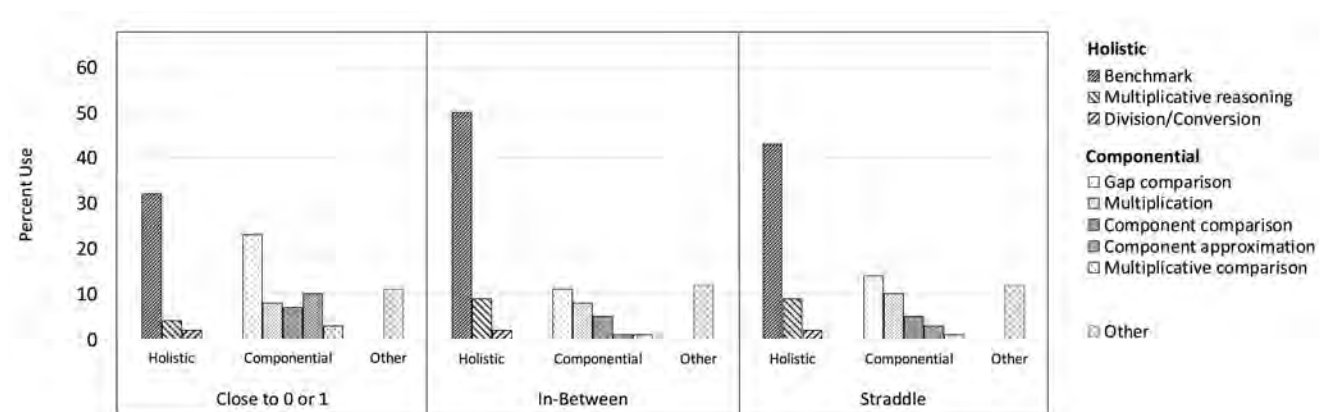
**Figure 1**

Percent Use of Holistic and Componential Strategies, for Each of the Three Item Types



**Figure 2**

Percent Use of Individual Holistic and Componential Strategies, for Each of the Three Item Types



Note. Percentages may not exactly correspond to those displayed in Figure 1 due to rounding.

Finally, we were interested in whether strategy use could partially explain the congruency effect. We had hypothesized that this could be the case if gap comparison was used more often or more efficiently for incongruent than congruent items. Table 5 displays strategy frequencies and efficiency scores for the two most frequent strategies, benchmarking and gap comparison, separately for congruent and incongruent items, and also displays differences in efficiency scores. While there were no substantial differences in frequencies of these strategies between congruent and incongruent items, it stands out that gap comparison was much more efficient on incongruent than congruent items, whereas the efficiency difference was much smaller for benchmarking. As noted earlier, gap comparison yields correct responses on all incongruent items, but not on all congruent items, which may explain the more efficient use of gap comparison on incongruent items overall. In our item set, gap comparison led to the correct response on only 10 of the 28 congruent items. When we removed the 18 congruent items for which gap comparison yielded an incorrect response, the efficiency score difference between congruent and incongruent items was reduced by 50% (from 11.8 to 5.9), but it did not fully disappear. Specifically, incongruent items were still solved significantly more accurately,  $M = 94\%$ ,  $SE = 1.1$ , than congruent ones,  $M = 86\%$ ,  $SE = 2.1$ , Wald  $\chi^2(1) = 13.58$ ,  $p < .001$ ,  $OR = 1.07$ , 95% CI [1.03, 1.13], although there was no longer a significant difference in response times,  $M = 4285$ ,  $SE = 175$  vs.  $M = 4424$ ,  $SE = 207$ ; Wald  $\chi^2(1) = 0.82$ ,  $p = .365$ , difference = 139, 95% CI [-164, 442]. In sum, the use of gap comparison explained some—but not all—of the congruency effect.

Table 5

Strategy Efficiency Scores for Congruent and Incongruent Items, for the Two Most Frequent Strategies

Strategy	Congruent		Incongruent		IES Difference (Cong–Incong)
	Percent Use	IES	Percent Use	IES	
Benchmarking	40	56.7	43	48.0	8.7
Gap comparison	14	67.9	18	37.4	30.6
Total		57.4		45.6	11.8

## Discussion

Adults can solve complex fraction comparisons quickly and accurately. On average, participants needed fewer than 5 seconds per item and solved 86% of these challenging comparisons correctly, in line with earlier research on challenging fraction comparisons in educated adults (Obersteiner et al., 2013, 2020; Schneider & Siegler, 2010). This generally high level of performance raises questions about what strategies adults use to compare fractions and how flexible and adaptive they are.

### Strategy Use in Fraction Comparison

Participants used holistic strategies, which rely on fraction magnitudes, far more frequently than componential strategies, which do not rely on fraction magnitudes. Thus, in line with earlier research with other methods (González-Forte et al., 2020; Obersteiner et al., 2013), our verbal report data suggest that the fraction magnitude comparison task with challenging items (but not with simple items) is a suitable means to assess fraction magnitude processing.

Our study allowed us to characterize the specific strategies that adults used to reason about fraction magnitudes. The most prominent holistic strategy involved using familiar numbers (especially 1 and  $\frac{1}{2}$ ) as benchmarks, in line with earlier research on simpler fraction comparisons (e.g., Clarke & Roche, 2009). The question remains how participants determined which benchmark was most suitable for any given item. For example, it is not clear why participants mentioned  $\frac{1}{3}$  as a benchmark more frequently than  $\frac{3}{4}$ . One explanation could be that  $\frac{1}{3}$  is a unit fraction (with numerator 1), which people may use more readily. On the other hand, one could argue that  $\frac{3}{4}$  is more familiar from everyday situations. Further research on item features that may encourage using specific benchmarks is needed.

Gap comparison was the second most frequent strategy overall, and the most frequent componential strategy. Although this strategy is not always valid, participants used it relatively frequently, as in previous studies of simpler fraction comparisons (e.g., Clarke & Roche, 2009; González-Forte et al., 2020). Our results showed that participants used gap comparison most frequently and very efficiently on close-to-0-or-1 items. This result reflects participants' adaptivity, because the subset of close-to-0-or-1 items included relatively more items for which gap comparison led to the correct response (31%, compared to 13% in each of the straddle and the in-between item subsets). Again, the question remains how participants decided whether gap comparison would be efficient for any given item. One explanation is that in the close-to-0-or-1 items, one fraction always had a particularly large or particularly small gap, and such gaps were presumably highly salient. It is also possible that participants sometimes used gap comparison first, and then reasoned about fractions' proximity to 0 or 1, which is a form of benchmarking. Although both strategies (gap comparison and benchmark) were assigned when participants mentioned them on the same trial, and we found combinations of gap comparison and benchmark very rarely, we cannot rule out the possibility that some participants used both strategies in combination but did not verbalize both.

We did not find a large effect of the tip that we gave some participants about the usefulness of benchmark strategies. This could mean that strategy use is not easily malleable. However, Fazio et al. (2016) found that university students' strategy use in fraction comparison was malleable. Given these past findings, we assume that the minimal effect of our tip was more likely due to the brief nature of the tip and to the high difficulty of our comparison items. We do not take our results as evidence against teaching approaches that encourage students to use various strategies. On the contrary,

we think that adaptive and flexible use of various strategies such as benchmarking should be a focus of mathematics instruction, as it may help participants build readily-accessible representations of fraction magnitudes.

## Natural Number Bias

Overall, participants displayed a reverse natural number bias. Our results suggest that this reverse bias may be associated with gap comparison more so than with simple comparisons between numerators or denominators, which people commonly employ when comparing fractions with common components (Fazio et al., 2016; Van Hoof et al., 2020). Our fraction comparison items did not have common components, and participants very rarely reported comparing components across fractions. Instead, people used gap comparison not only on incongruent items, for which it was very efficient, but also, albeit somewhat less frequently, on congruent items, for which it was very inefficient. Participants' limited adaptivity in using gap comparison could partly explain the reverse bias pattern.

What other factors could contribute to the reverse bias pattern? One possibility is subtle processing mechanisms that are not accessible to verbal strategy reports (e.g., intuitive processes). Although some studies have identified intuitive processes, such as automatic activation of natural number magnitudes, as contributing to the *typical* bias pattern (Van Hoof et al., 2020), these processes cannot explain the *reverse* bias pattern. It may be that item congruency is related to other salient features that have not yet been identified.

## Limitations

One limitation of this work concerns the validity of retrospective self-reports of strategy use. Our experimental procedure was in line with accepted criteria for valid self-reports (Ericsson & Simon, 1993; i.e., solution time longer than 1 sec, response within 10 seconds). However, some participants struggled to explain their strategies (about 5% of responses were missing or unintelligible), and prompting participants to report their strategies may have affected their strategy choices. Even so, the accuracy and response time data were largely in line with previous work that used the same items but did not require verbal reports (Obersteiner et al., 2020), and most of the strategies we identified have been previously documented (see, e.g., Fazio et al., 2016).

A second limitation is that this work does not provide definitive evidence about the sequential and temporal processes involved in fraction comparison. For example, people might start out by considering differences (rather than ratios) between numerators and denominators, and if these differences provide enough information for them to feel confident in their magnitude comparisons, they may base their responses on these differences (i.e., using the gap comparison strategy). If, however, gap comparison is inconclusive, or if other item features encourage holistic reasoning, they may then engage in more effortful holistic reasoning. Some adults, knowing that gap comparison is not a universally applicable heuristic, may bypass such reasoning altogether, and engage immediately in holistic reasoning. Further work is needed to delineate the processes involved in selecting and applying comparison strategies.

## Conclusion

In sum, this study revealed that educated adults often rely on fraction magnitudes when comparing challenging fraction pairs. Thus, the fraction comparison task with challenging items is suitable for assessing fraction magnitude processing. The present work highlights both flexibility and adaptivity in adults' strategies for fraction comparison, and it demonstrates that strategy choices contribute to the occurrence of the reverse natural number bias.



**Funding:** This work was supported by a Feodor Lynen Research Fellowship granted to Andreas Obersteiner by the Alexander von Humboldt Foundation, Germany.

**Acknowledgments:** The authors have no additional (i.e., non-financial) support to report.

**Competing Interests:** The first author is an Associate Editor of the *Journal of Numerical Cognition* but played no editorial role in this particular article and did not intervene in any form in the peer review process. All authors declare that no competing interests exist.

**Ethics Statement:** This study has been carried out in accordance with ethical and scientific standards. The study was approved by the Institutional Review Board at the University of Wisconsin–Madison, United States.

## Supplementary Materials

The Supplementary Materials contain one pdf file with separate analyses of strategy data for participants in the two experimental groups “With Tip” and “Without Tip”. Specifically, the file includes separate analyses of Tables 3, 4, and 5, and Figures 1 and 2 (for access see [Index of Supplementary Materials](#) below).

### Index of Supplementary Materials

Obersteiner, A., Alibali, M. W., & Marupudi, V. (2022). *Supplementary materials to "Comparing fraction magnitudes: Adults' verbal reports reveal strategy flexibility and adaptivity, but also bias"* [Additional analyses]. PsychOpen GOLD. <https://doi.org/10.23668/psycharchives.8223>

## References

- Alibali, M. W., & Sidney, P. G. (2015). Variability in the natural number bias: Who, when, how, and why. *Learning and Instruction, 37*, 56–61. <https://doi.org/10.1016/j.learninstruc.2015.01.003>
- Barraza, P., Avaria, R., & Leiva, I. (2017). The role of attentional networks in the access to the numerical magnitude of fractions in adults / El rol de las redes atencionales en el acceso a la magnitud numérica de fracciones en adultos. *Estudios de Psicología, 38*(2), 495–522. <https://doi.org/10.1080/02109395.2017.1295575>
- 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), 1410–1419. <https://doi.org/10.1037/0096-1523.33.6.1410>
- Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in Cognitive Psychology: Is the Inverse Efficiency Score (IES) a better dependent variable than the mean Reaction Time (RT) and the Percentage Of Errors (PE)? *Psychologica Belgica, 51*(1), 5–13. <https://doi.org/10.5334/pb-51-1-5>
- Clarke, D. M., & Roche, A. (2009). Students' fraction comparison strategies as a window into robust understanding and possible pointers for instruction. *Educational Studies in Mathematics, 72*(1), 127–138. <https://doi.org/10.1007/s10649-009-9198-9>
- DeWolf, M., & Vosniadou, S. (2015). The representation of fraction magnitudes and the whole number bias reconsidered. *Learning and Instruction, 37*, 39–49. <https://doi.org/10.1016/j.learninstruc.2014.07.002>
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA, USA: MIT Press.
- Faulkenberry, T. J., & Pierce, B. H. (2011). Mental representations in fraction comparison. *Experimental Psychology, 58*(6), 480–489. <https://doi.org/10.1027/1618-3169/a000116>
- Fazio, L. K., DeWolf, M., & Siegler, R. S. (2016). Strategy use and strategy choice in fraction magnitude comparison. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(1), 1–16. <https://doi.org/10.1037/xlm0000153>
- Fischbein, E. (1987). *Intuition in science and mathematics: An educational approach*. Dordrecht, The Netherlands: Reidel.
- Ganor-Stern, D., Karasik-Rivkin, I., & Tzelgov, J. (2011). Holistic representation of unit fractions. *Experimental Psychology, 58*(3), 201–206. <https://doi.org/10.1027/1618-3169/a000086>
- Gillard, E., Van Dooren, W., Schaeken, W., & Verschaffel, L. (2009). Dual processes in the psychology of mathematics education and cognitive psychology. *Human Development, 52*, 95–108. <https://doi.org/10.1159/000202728>

- Gómez, D. M., & Dartnell, P. (2019). Middle schoolers' biases and strategies in a fraction comparison task. *International Journal of Science and Mathematics Education*, 17(6), 1233–1250. <https://doi.org/10.1007/s10763-018-9913-z>
- González-Forte, J. M., Fernández, C., Van Hoof, J., & Van Dooren, W. (2019). Exploring students' reasoning about fraction magnitude. In M. Graven, H. Venkat, A. Essien, & P. Vale (Eds.), *Proceedings of the 43rd PME Conference* (Vol. 2, pp. 272–279). Pretoria, South Africa: PME.
- González-Forte, J. M., Fernández, C., Van Hoof, J., & Van Dooren, W. (2020). Various ways to determine rational number size: An exploration across primary and secondary education. *European Journal of Psychology of Education*, 35(3), 549–565. <https://doi.org/10.1007/s10212-019-00440-w>
- Heinze, A., Star, J. R., & Verschaffel, L. (2009). Flexible and adaptive use of strategies and representations in mathematics education. *ZDM*, 41(5), 535–540. <https://doi.org/10.1007/s11858-009-0214-4>
- Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General*, 124(1), 83–97. <https://doi.org/10.1037/0096-3445.124.1.83>
- Liu, F. (2018). Mental representation of fractions: It all depends on whether they are common or uncommon. *Quarterly Journal of Experimental Psychology*, 71(9), 1873–1886. <https://doi.org/10.1080/17470218.2017.1366532>
- 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. <https://doi.org/10.1016/j.actpsy.2010.07.014>
- Morales, N., Dartnell, P., & Gómez, D. M. (2020). A study on congruency effects and numerical distance in fraction comparison by expert undergraduate students. *Frontiers in Psychology*, 11, Article 1190. <https://doi.org/10.3389/fpsyg.2020.01190>
- Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, 215, 1519–1520. <https://doi.org/10.1038/2151519a0>
- 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. [https://doi.org/10.1207/s15326985ep4001\\_3](https://doi.org/10.1207/s15326985ep4001_3)
- Obersteiner, A., Alibali, M. W., & Marupudi, V. (2020). Complex fraction comparisons and the natural number bias: The role of benchmarks. *Learning and Instruction*, 67, Article 101307. <https://doi.org/10.1016/j.learninstruc.2020.101307>
- Obersteiner, A., Marupudi, V., & Alibali, M. W. (2019). Adults' strategy use in fraction comparison. In M. Graven, H. Venkat, A. Essien, & P. Vale (Eds.), *Proceedings of the 43rd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 153–160). Pretoria, South Africa: PME.
- Obersteiner, A., & Tumpek, C. (2016). Measuring fraction comparison strategies with eye-tracking. *ZDM Mathematics Education*, 48(3), 255–266. <https://doi.org/10.1007/s11858-015-0742-z>
- Obersteiner, A., Van Dooren, W., Van Hoof, J., & Verschaffel, L. (2013). The natural number bias and magnitude representation in fraction comparison by expert mathematicians. *Learning and Instruction*, 28, 64–72. <https://doi.org/10.1016/j.learninstruc.2013.05.003>
- Reinhold, F., Obersteiner, A., Hoch, S., Hofer, S. I., & Reiss, K. (2020). The interplay between the natural number bias and fraction magnitude processing in low-achieving students. *Frontiers in Education*, 5, Article 29. <https://doi.org/10.3389/feduc.2020.00029>
- Rinne, L. F., Ye, A., & Jordan, N. C. (2017). Development of fraction comparison strategies: A latent transition analysis. *Developmental Psychology*, 53(4), 713–730. <https://doi.org/10.1037/dev0000275>
- Schneider, M., & Siegler, R. S. (2010). Representations of the magnitudes of fractions. *Journal of Experimental Psychology: Human Perception and Performance*, 36(5), 1227–1238. <https://doi.org/10.1037/a0018170>
- Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperregu, I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, 23(7), 691–697. <https://doi.org/10.1177/0956797612440101>
- Siegler, R. S., & Lemaire, P. (1997). Older and younger adults' strategy choices in multiplication: Testing predictions of ASCM using the choice/no-choice method. *Journal of Experimental Psychology: General*, 126(1), 71–92. <https://doi.org/10.1037/0096-3445.126.1.71>
- Stafylidou, S., & Vosniadou, S. (2004). The development of students' understanding of the numerical value of fractions. *Learning and Instruction*, 14(5), 503–518. <https://doi.org/10.1016/j.learninstruc.2004.06.015>
- 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. <https://doi.org/10.1080/14794802.2013.797747>

Van Hoof, J., Verschaffel, L., De Neys, W., & Van Dooren, W. (2020). Intuitive errors in learners' fraction understanding: A dual-process perspective on the natural number bias. *Memory & Cognition*, 48(7), 1171–1180. <https://doi.org/10.3758/s13421-020-01045-1>

## Appendix

**Table A.1**

*List of Fraction Comparison Items*

Congruent	Incongruent
<b>Straddle</b>	
6/37 vs. 13/43	17/91 vs. 13/41
5/27 vs. 28/87	17/93 vs. 5/16
26/81 vs. 4/25	6/19 vs. 19/96
31/98 vs. 5/31	13/43 vs. 14/97
11/26 vs. 34/59	31/71 vs. 13/23
13/31 vs. 43/75	23/52 vs. 11/19
32/57 vs. 11/27	16/27 vs. 28/65
41/73 vs. 15/34	13/22 vs. 27/61
13/19 vs. 25/31	26/37 vs. 19/22
16/23 vs. 35/41	49/71 vs. 35/43
56/67 vs. 15/22	11/13 vs. 16/23
47/56 vs. 29/41	23/27 vs. 27/38
<b>In-Between</b>	
11/37 vs. 42/95	19/63 vs. 16/35
11/36 vs. 41/92	18/61 vs. 8/19
22/49 vs. 13/41	11/25 vs. 27/91
31/69 vs. 9/29	13/29 vs. 29/95
21/37 vs. 41/59	17/31 vs. 11/16
11/19 vs. 25/36	27/49 vs. 19/27
37/52 vs. 14/25	22/31 vs. 37/64
19/27 vs. 13/23	23/33 vs. 27/47
<b>Close to 0 or 1</b>	
2/21 vs. 21/97	10/99 vs. 9/43
3/49 vs. 19/94	6/97 vs. 4/24
19/97 vs. 5/96	6/29 vs. 7/99
13/63 vs. 3/58	9/44 vs. 11/98
21/26 vs. 59/62	68/83 vs. 49/52
31/38 vs. 55/58	30/37 vs. 27/29
41/43 vs. 17/21	46/49 vs. 53/66
58/61 vs. 13/16	35/37 vs. 49/61



*Journal of Numerical Cognition* (JNC) is an official journal of the Mathematical Cognition and Learning Society (MCLS).



leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.