Journal of Experimental Child Psychology 263 (2026) 106373

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

# Characterizing error types in the comprehension of fractions: The number line test

Maxime Cauté [a,*], Cassandra Potier Watkins [a,b], Chenxi He [a], Stanislas Dehaene [a,b,*]

[a] Cognitive Neuroimaging Unit (CEA, INSERM, Université Paris-Saclay), Centre d'études de Saclay, Bâtiment 145, 91191 Gif-sur-Yvette, France
[b] Collège de France (Université Paris-Sciences-Lettres (PSL)), 11 place Marcelin-Berthelot, 75231 Paris Cedex 05, France

ARTICLE INFO

*Keywords:*
Number line
Fractions
Error type
Strategy choice
Partition
Conversion

ABSTRACT

Understanding fractions is a major hurdle for many students. A key aspect of fraction comprehension is the ability to evaluate their numerical magnitude. Here, we use a number-to-line task, where students point to the location of a number on a graduated line, to characterize errors in fraction comprehension. A total of ∼ 26,000 French pupils from 6th to 10th grade were tested (U. S. equivalent grades). Error rates were high, almost 80 % in 6th grade and 45 % in 10th grade. Errors could be classified into seven dominant patterns, whose frequency varied by grade level and individual performance. Younger and lower-performing children mostly confused fractions with decimals. Older and higher-performing children often confused a fraction $\frac{a}{b}$ and its inverse $\frac{b}{a}$. All grades also confused the roles of the numerator and the denominator. We propose a theoretical framework suggesting that errors arise as bugs in the execution of one of two main strategies: children converting the fraction into a decimal, or partitioning the line into units and counting them. This model explains the observed error patterns as stemming from inappropriate strategy selection, flawed execution, or incorrect corrective steps due to flawed execution. Our analysis provides a deeper understanding of the various traps that students face when interpreting a fraction's magnitude, the frequency of these errors, and their sequential order.

## Introduction

Fractions are a challenge for math learners (Ni & Zhou, 2005; Siegler, Fazio, Bailey, & Zhou, 2013). Even after years of practice, many children continue to fail even some of the most basic tests. For instance, Stafylidou and Vosniadou (2004) reported that, by 9th grade, 42 % of Greek children could not correctly order the fractions $\frac{1}{7}$, $\frac{5}{6}$, 1 and $\frac{4}{3}$. In fraction arithmetic, Siegler and Pyke (2013) reported error rates as high as 43 % for US 8th graders. Even more striking, Hannula (2003) reported that 50 % of Finnish 7th graders could not place ¾ between 0 and 1 when prompted to place it on a number line. These difficulties are often attributed to a lack of conceptual understanding of fractions, or at the very least, a superficial grasp of the subject (Ni & Zhou, 2005; Stafylidou & Vosniadou, 2004; Chan, Leu, & Chen, 2007). Indeed, children may successfully solve school-based tests by mere rote memorization of the specific procedure that was taught, without any deep understanding. Braithwaite and Siegler (Braithwaite, Pyke, & Siegler, 2017; Braithwaite & Siegler, 2023) even proposed that children's behavior in arithmetic could be replicated by a purely procedural model trained on exercises found in textbooks. As a result, determining whether a child has a true understanding of the meaning of fractions remains a

* Corresponding authors at: Collège de France (Université Paris-Sciences-Lettres (PSL)), 11 place Marcelin-Berthelot, 75231 Paris Cedex 05, France

*E-mail addresses:* maxime.caute@college-de-france.fr (M. Cauté), stanislas.dehaene@college-de-france.fr (S. Dehaene).

https://doi.org/10.1016/j.jecp.2025.106373
Received 20 February 2025; Received in revised form 14 August 2025;
Available online 12 November 2025

challenge.

Some researchers have proposed distinguishing several different meanings or 'subconstructs' for fractions (Kieren, 1976; Behr, Lesh, Post, & Silver, 1983). Among these, converging evidence suggests that magnitude, which describes how large a fraction is, plays a central role in understanding fractions (Siegler, Thompson, & Schneider, 2011). Decades of research indicate that magnitude is, in fact, a privileged representation of numbers in the brain. The magnitude dimension of whole numbers is typically processed by the Approximate Number System (ANS), located in the intraparietal sulcus (IPS) (Dehaene, Piazza, Pinel, & Cohen, 2003), which encodes the approximate magnitude of numbers presented in both symbolic (e.g., Arabic numerals) and non-symbolic numbers (e.g., sets of dots). A major signature of the ANS is the ratio-sensitive distance effect, whereby two numbers are easier to discriminate when their ratio is greater (Moyer & Landauer, 1967). Some studies have reported a similar pattern in tasks involving both symbolic and non-symbolic displays of fractions (Kallai & Tzelgov, 2009; Gabriel, Szucs, & Content, 2013; Matthews & Chesney, 2015). Crucially, the distance between components (i.e., the numerator or the denominator alone) did not suffice to explain the effect, and some authors therefore postulated that the ANS (Gallistel & Gelman, 2000) or another similar system (Lewis, Matthews, & Hubbard, 2016; Matthews & Chesney, 2015) may encode the magnitude of fractions. fMRI studies on symbolic fractions have found further evidence for the encoding of magnitudes in and around the IPS, close to or overlapping with the location of activations in tasks soliciting the ANS (Ischebeck, Schocke, & Delazer, 2009; Jacob & Nieder, 2009). As an example, Jacob and Nieder (2009) habituated participants to equivalent symbolic fractions of a given magnitude (⅙), and observed that fMRI activation in the anterior IPS increased in response to deviant fractions of a different magnitude. This occurred even in the absence of a task, and the (linear) distance between the two magnitudes further modulated the activation.

It remains uncertain whether magnitude is systematically and automatically encoded for symbolic fractions (Bhatia et al., 2020; Gabriel et al., 2013; Kallai & Tzelgov, 2009). Nevertheless, accuracy in estimating the magnitude of non-symbolic fractions predicts performance with symbolic fractions (Matthews, Lewis, & Hubbard, 2016). Also, training children to estimate fraction magnitudes has often proven effective in improving their overall performance with fractions, with demonstrated transfer to new tasks (Fazio, Kennedy, & Siegler, 2016; Hamdan & Gunderson, 2017). However, some studies have failed to observe such transfer, or even any improvement on the training task (Tian, Bartek, Rahman, & Gunderson, 2021; Nuraydin, Stricker, & Schneider, 2022). To quote Siegler et al. (2013), "[a] common feature of the successful interventions is that they help children understand how symbolic fractions map onto the magnitudes they represent."

The metaphor of the number line provides a straightforward way to explain and to represent fraction magnitudes. Each magnitude can be assigned a position on the line, and since this is true of any real number, it also helps to integrate fractions with whole numbers (Siegler et al., 2011). Accordingly, tasks requiring to point to the number line have been extensively used to investigate number comprehension (Siegler & Opfer, 2003; Dotan & Dehaene, 2013; Pinheiro-Chagas, Dotan, Piazza, & Dehaene, 2017; Obersteiner & Hofreiter, 2017; Patel & Varma, 2018), including the comprehension of fractions (Behr et al., 1983; Siegler & Thompson, 2014). Number line tasks tend to prove more difficult, and therefore more revealing of underlying comprehension difficulties, than other tasks (Behr et al., 1983; Hannula, 2003). For instance, within a sample of 77 4th graders, virtually all of them could shade ⅔ or ¾ of a rectangle or circle (already divided in thirds or fourths), but only about a third could place these fractions on a number line, even though it too was segmented (Behr et al., 1983). The placement of graduations appears to be a major factor in children's performance (Siegler & Thompson, 2014). In particular, children make many more errors when the graduations do not match the denominator of the fraction (Larson, 1980; Behr et al., 1983; Bright, Behr, Post, & Wachsmuth, 1988). They are also easily confused when the boundaries of the line do not align with the 0–1 segment (Larson, 1980; Bright et al., 1988).

These observations underscore the specific challenges that certain aspects of fraction knowledge present to learners. However, to our knowledge, no study has attempted a systematic analysis of their errors. Hannula (2003) asked 7th graders to place ¾ on a line intersected with marks for 0 and 1, and noted that some students confused that fraction with 3.4, counted 3 times the 0–1 interval (thus answering 3), or placed it at ¾ of the whole line instead of the 0–1 segment. However, these were just illustrative examples, and participants were only asked about the fraction ¾, which limits the generalizability of these findings. Bright et al. (1988) tested children on number lines to evaluate the effectiveness of an intervention. They observed that some errors could be explained by children using the entire line as the unit (42 % at pretest; 32 % at posttest), by counting marks instead of intervals to know in how many pieces the unit is split (16 %; 0 %), or by pointing to the inverse of the target fraction (16 %; 24 %). Again, however, the sample size was limited, with only eight children.

Building upon this background literature, the goal of the present work is to characterize the errors made by a representative sample of French students when placing fractions on a number line, and to assess to what extent this task could diagnose specific gaps in their understanding of the magnitude of fractions. To this end, we took advantage of France's yearly national evaluations (Groupe de travail du Csen. (2019), 2019). In collaboration with the department of evaluation of the French ministry of education (DEPP), we obtained detailed data from a computerized number-line pointing task in ~26,000 students from 6th to 10th grade (U.S. grade equivalents). Students were presented with a graduated number line and prompted to place target numbers, including five fractions, at their exact positions on the line. All fractions corresponded to an existing mark on the line and subjects could only select one of these marks, allowing us to accurately determine the intended response. This task, with our chosen set of targets, was part of the national French curriculum in 4th grade, which states that fractions should be studied in 4th and 5th grades and be consolidated by 6th grade. In spite of these explicit instructions, and contrary to grade expectations, we instead observed high rates of errors across all grades, revealing several common error patterns: confusion with decimals (Hannula, 2003) and with mixed numbers, miscount of graduations (Bright et al., 1988; Hannula, 2003), misidentification of the line as the unit (Bright et al., 1988; Hannula, 2003), and inversion of the target fraction (Bright et al., 1988). We develop a model for the origin of these errors and how they evolve across grades.

## Study 1: Number line in French 6th graders

In experiment 1, we introduced our computerized number-line pointing task and described the errors made by a sample of 5707 French 6th graders, leveraging France’s annual national evaluations.

### *Material & methods*

#### *Procedure*

The experiment was run on computers already available in schools, with students using a mouse to respond. Target numbers were shown one at a time in standard mathematical notation and displayed on a yellow rectangle at the center of the screen, with a pin attached to the bottom, resembling a sign (Fig. 1). Students had 10 s to answer by dragging the sign to a position on a graduated number line at the bottom of the screen, aligning the pin with their estimated location for the number on the line. Once dropped, the sign snapped to the closest graduation, providing a small margin of error (pilot studies showed that it improved participants’ overall accuracy). Students could not move their response after it was snapped. Immediate feedback was given, indicating whether the response was correct or incorrect, and showing the correct response for the latter.

Participants completed two randomly-ordered blocks of trials. Each block featured a different line, either graduated from 0 to 5 or from 0 to 20. Both lines featured three levels of graduations of decreasing height and thickness: major, intermediate and minor. The major graduations were labeled with their numerical values above them. The 0–20 line was used for testing whole numbers and their arithmetic (e.g., place $6 + 7$), and will not be considered any further in the analysis, as we focus here exclusively on fractions. The other line, ranging from 0 to 5, was decimally graduated and is depicted in Fig. 1. Students were familiarized with the lines at the start of each block using a single training trial in which they were asked to place the number ‘5′. Feedback was given for the trial. The data from this trial was not considered in the analyses.

#### *Stimuli*

The target stimuli were selected with two constraints in mind. The experiment had to assess children’s proficiency with a variety of numbers and operations (although we only report fractions here), but sessions had to remain relatively short to reduce fatigue, especially since children were taking several other tests during the national evaluations period. Thus, 10 partially overlapping lists of target items were prepared, and each participant was randomly assigned to a list. The lists covered 11 categories in total, with five targets per category in each list, resulting in 55 test trials per participant. The first five categories only involved whole numbers and operations used with the 0–20 line block. Five whole numbers were presented during this block and served as a control for participant’s comprehension of the task. The remaining six categories of targets appeared during the 0–5 line block in a random order. One category only involved whole numbers, just as for the 0–20 line, and again served as a control to ensure task comprehension. The other categories revolved around rational numbers: decimal numbers (written with a comma as per traditional French notation, e.g., ‘3,6′); decimal system operations (a whole number plus a one-digit decimal value, e.g., ‘3 + 0.6′ or ‘0.6 + 3′); calculation with decimals (addition/subtraction of decimals whose decimal value was non-zero single-digit, e.g., ‘2.6–1.5′); fractions; and calculation with fractions (additions/subtractions of a fraction with either a same-denominator fraction or a whole number). All fractions were written vertically, e.g., $\frac{3}{6}$. The full list of stimuli is reported in supplementary Table S1.

There were 21 unique fractions across the 10 lists, split in four categories. The first category covered fractions in the format $\frac{n}{1}$ ($\frac{2}{1}$; $\frac{3}{1}$; $\frac{4}{1}$; $\frac{5}{1}$). The second category comprised other fractions equal to a whole number ($\frac{4}{2}$; $\frac{6}{2}$; $\frac{8}{2}$; $\frac{10}{2}$; $\frac{6}{3}$; $\frac{9}{3}$; $\frac{8}{4}$; $\frac{10}{5}$). The third category comprised fractions with tenths ($\frac{1}{10}$; $\frac{2}{10}$; $\frac{3}{10}$; $\frac{4}{10}$; $\frac{5}{10}$), and the last comprised fractions equivalent to one half ($\frac{1}{2}$; $\frac{2}{4}$; $\frac{3}{6}$; $\frac{4}{8}$). Note that $\frac{5}{10}$ could also fit into the category of equivalents to one half. We arbitrarily classified it as a fraction expressed in tenths to obtain non-overlapping sets for our analyses, and because the denominator 10 was arguably more salient than the equivalence to one half.

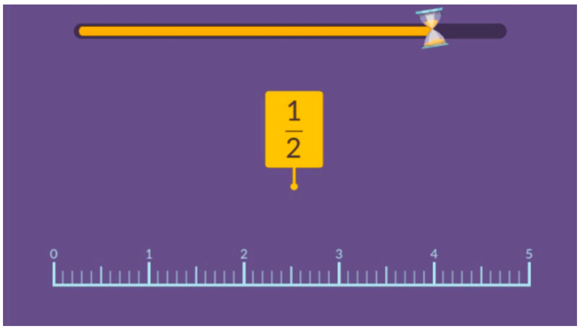


**Fig. 1. Screenshot of the computerized number-to-line task**. On each trial the child must, within 10 s, click on the precise location for a target number on a number line graduated from 0 to 5.

*Participants*

7800 6th graders took the test at school in September 2022, at the beginning of the school year. This sample was selected to be representative of French 6th graders, who, in France, enter this grade the (calendar) year they turn 11. Children in our sample were aged 11 years and 4 months on average. The data collection occurred in France where data on race/ethnicity cannot legally be collected. We only kept children who undoubtedly understood the number-to-line test. Thus, we excluded children who were part of a supplementary group who did not receive feedback, and those who made two errors or more on integers on either line (error rate > 20 %). The final sample and detailed figures per exclusion criterion are reported in Table 1. Post-hoc analyses on the subset of participants who did not receive feedback (again restricted to those who correctly placed the integers) revealed that their distribution of responses was highly correlated to that of participants who did receive feedback (Pearson's $\rho = 0.95$).

*Transparency and openness statement*

The methods and analyses for this study were not pre-registered. We report all key measures assessed during the study, with the exception of reaction times (RT), given our focus on error types and the large time span allowed for responses (10 s). The data used in this study is the property of the French DEPP (Department of Evaluation, Prospective, and Performance) and may not be made publicly available without their consent. However, analytic scripts are provided on the following Open Science Framework (OSF): https://osf.io/267s5/.

## Results

We could confirm that our procedure was robust enough to analyze children's errors on fractions by looking at data from whole numbers and decimal targets. Most children could place these numbers correctly (a child made on average 4 % errors on whole numbers, 13 % on decimals). Since children were assigned to groups with different lists of stimuli, we also confirmed that all groups responded similarly. Each of the 21 fractions was presented to at least two groups, and the average correlation between responses was as high as 0.97. Consequently, we merged all the responses from all groups together in the following analyses. Unless specified otherwise, analyses over multiple fractions were also always computed after averaging across children for each fraction. While this reduces the total number of data points in our statistics, it allows us to adjust our statistical power to the fact that children were only tested on a limited set of 21 fractions.

*Overall performance*

Children's accuracy on fractions was very low: on average, each child made 79 % of errors on its set of target fractions. Error rates on individual fractions ranged from 62 % (on $\frac{5}{1}$) to 93 % (on $\frac{3}{6}$). The full distribution of error rates for each target is displayed in the left panel of Fig. 2. Some targets yielded many more errors than others, as confirmed by a two-way chi-square test over correct vs. incorrect responses by target ($\chi^2(20) = 1341.04$, $p < .001$). Surprisingly, $\frac{1}{2}$ was not particularly easy for children (average error rate of 78 %) despite being a very familiar fraction of the set. Error rates also differed between fraction categories: fractions equivalent to one half yielded the most errors (on average, fractions in this category yielded 88 % of errors), followed by fractions with a whole number value (excluding fractions of the form $\frac{n}{1}$; 82 %), then fractions of the form $\frac{n}{1}$ (73 %), or $\frac{n}{10}$ (72 %). Significance was confirmed by performing an ANOVA over error rates per target ($F(3, 17) = 7.69$, $p = .002$, $\eta^2 = 0.58$).

Because each participant only judged five fractions, average error rates could be partially affected by variations in participant proficiency. To separate this confound, we estimated each target's difficulty using Item Response Theory (IRT) (Cai, Choi, Hansen, & Harrell, 2016). The results are displayed on the right panel of Fig. 2, and are significantly correlated with the observed error rates ($F(1, 19) = 92.25$, $R^2 = .83$; $\beta = 0.05$, $t(19) = 9.60$, $p < .001$). More details about our IRT analysis can be found in supplementary materials.

*Errors*

The specific errors made by children were also informative. For each target, a chi-square test across positions revealed that errors were not uniformly distributed (min $\chi^2(49) = 3595.43$, all $ps < .001$; p-values were FDR-corrected for multiple comparisons across

**Table 1**
**Characteristics of our participant sample.** Our initial sample was selected to be representative within each grade and curriculum. We excluded a number of participants from the analysis because of technical issues or uncertain comprehension of the task.

| | 6th grade | 8th grade | 10th grade | | |
|---|---|---|---|---|---|
| | | | General | Professional | Applied |
| Initial Sample | 7800 | 5650 | 7710 | 7205 | 5393 |
| Mean age(month; age) | 11;4 | 13;4 | 15;3 | 15;6 | 15;11 |
| Unreadable data | 3 | – | – | – | 1 |
| Lack of responses on one line | 99 | 0 | 114 | 20 | 31 |
| Lack of feedback (supplementary group) | 1316 | 579 | 1282 | 1231 | 479 |
| Failure on integer targets | 675 | 541 | 238 | 437 | 688 |
| Final sample | 5707 | 4530 | 6076 | 5517 | 4194 |
| Incl. girls | 2793 | 2224 | 3226 | 2369 | 1672 |

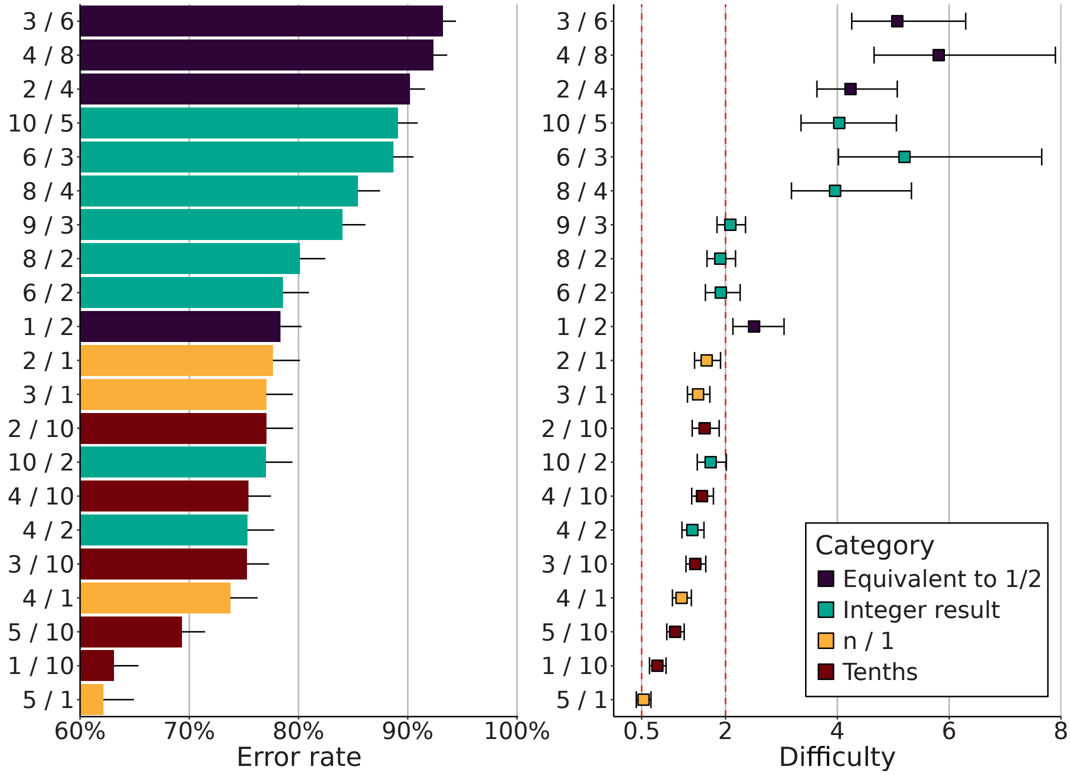


**Fig. 2. Ranking of the fractions used in our study according to their error rate in 6th graders (left).** The dots on the right indicate their difficulty as assessed by item response theory (IRT). We only used very simple fractions which could be classified into four subcategories: those equal or equivalent to $\frac{1}{2}$ (e.g., $\frac{3}{6}$), those involving division by 1 (e.g., $\frac{4}{1}$), those yielding an integer in another manner (e.g., $\frac{6}{2}$), and those involving division by 10 (see color legends). For IRT difficulty (right), dashed red bars correspond to boundaries between levels classically considered as normal, hard (0.5) and very hard (2). On the left panel, error bars indicate one standard error of the mean. On the right panel, they indicate 95% CI interval from IRT estimation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

targets). This result held even after removing integers (min $\chi^2(43) = 584.42$; all $ps < .001$), indicating that it could not be explained by integers serving as attractors. Overall, responses were focused on a limited set of positions, with half of the errors for a given target falling, on average, at just 3.05 specific positions. We also confirmed that most errors were not simply driven by difficulty in placing the fraction accurately at the position desired. If so, the errors would have predominantly been produced at a location adjacent to the correct response, but such errors were exceedingly rare: the median proportion of responses falling one graduation before the correct response was 0.76 %, computed across all stimuli; and this figure was 0.75 % for responses falling one graduation after the correct response (median computed across all 19 stimuli whose response was not 5, which was the end of the line).

For illustration, Fig. 3 shows the distributions of children's responses to four target fractions: $\frac{1}{2}$ and $\frac{3}{6}$ on the first row (both equivalent to 0.5), and the inverse fractions $\frac{2}{1}$ and $\frac{6}{3}$ on the second (both equivalent to 2). If children understood fraction equivalence, one would expect distributions to be similar between equivalent fractions, displayed on the same row. We observe the contrary: distributions are much more similar across columns, that is across inverse fractions. To test this observation, we considered the 6 pairs of target fractions where one is the inverse of the other: $\frac{1}{2}$ / $\frac{2}{1}$, $\frac{2}{4}$ / $\frac{4}{2}$, $\frac{3}{6}$ / $\frac{6}{3}$, $\frac{4}{8}$ / $\frac{8}{4}$, $\frac{2}{10}$ / $\frac{10}{2}$, and $\frac{5}{10}$ / $\frac{10}{5}$. Within and across the pairs, we computed how much the distributions were pairwise correlated using Pearson's $r$ (Fig. 4). As shown in Fig. 4, pairs of inverted fractions led to significantly correlated response distributions (average $r = .88$), much more than other pairs (average $r = .25$). An exact permutation test found that the six pairs of inverse fractions were the second best pairing of 90,858,768 possibilities ($p < .001$). Thus, children tended to respond identically to fractions $\frac{a}{b}$ and $\frac{b}{a}$, suggesting that many did not understand or pay enough attention to the distinct roles played by the numerator and the denominator.

Going back to the response distributions displayed in Fig. 2, we also observed that the most frequent response positions corresponded to consistent patterns, colored on the figure. We now introduce a terminology for the various response patterns. Two patterns were related to the magnitude of the fraction: one was the *correct* position (e.g., responding 0.5 for $\frac{1}{2}$ or $\frac{3}{6}$), while the other was *correct relative to the line*, i.e. it was the correct position for the fraction multiplied by a factor of 5, corresponding to the length of the 0–5 number line (e.g., responding 2.5 for $\frac{1}{2}$ or $\frac{3}{6}$; this error is hereafter termed *relative*). Although the *relative* error pattern makes valid predictions only for fractions smaller than 1, we will elaborate later on the existence of an inverted *relative* error pattern that can apply to fractions greater than 1. Participants also occasionally selected the inverse of the target fraction (e.g., responding 2 for $\frac{3}{6}$). We called

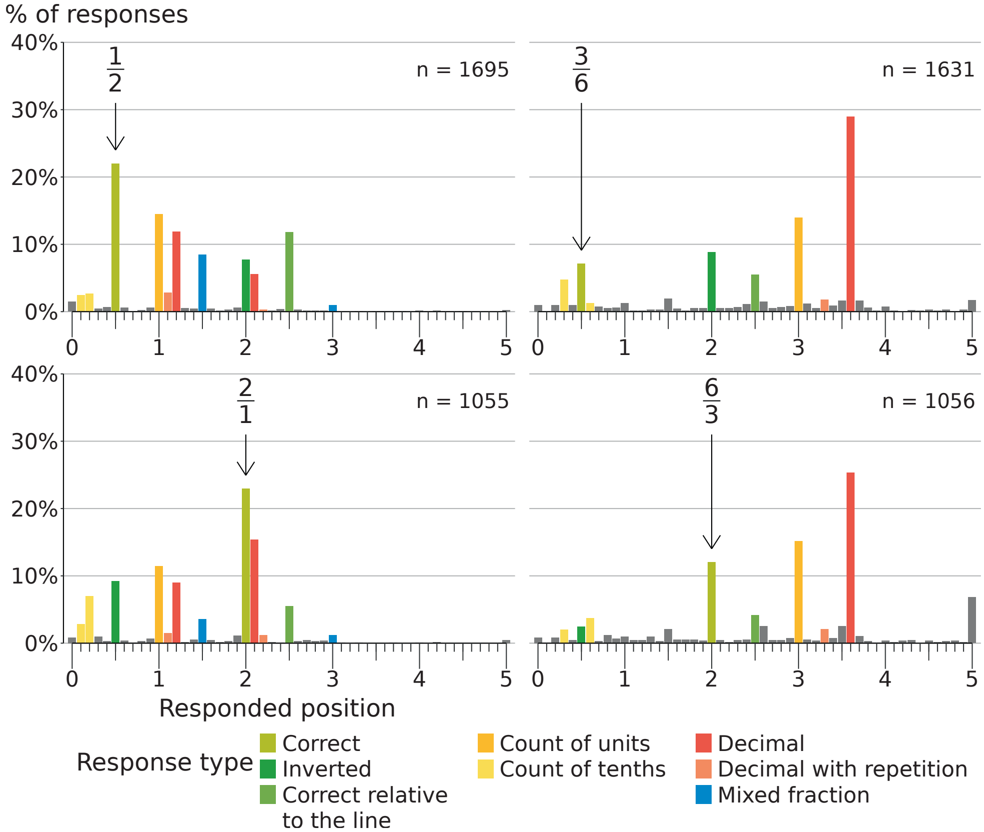


**Fig. 3. Examples of response distributions for four target fractions:** $\frac{1}{2}$, $\frac{3}{6}$, $\frac{2}{1}$ **and** $\frac{6}{3}$**.** Fractions on the same row have equivalent magnitudes, while fractions on the same column are the inverse of one another. The correct response is marked by an arrow. Colors correspond to various types of errors, as indicated in the legend.

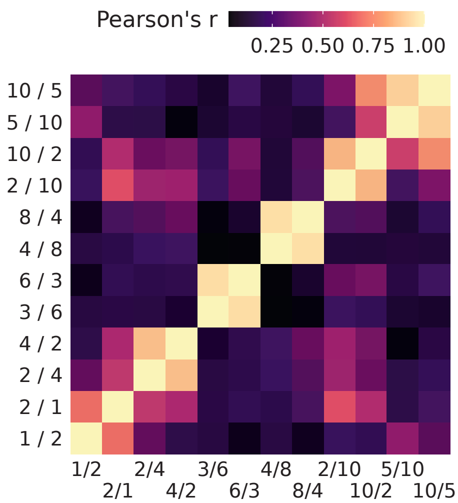


**Fig. 4. Correlation between response distributions for fractions whose inverse was also present in the target list.** Note the high correlation of responses to fractions and to their inverses, thus indicating that participants were confused about the roles of the numerator and denominator.

this the *inversion* pattern.

Two other response patterns matched a single component of the fraction, taken either as units (*count of units*; e.g., responding 1 or 2 for $\frac{1}{2}$) or as tenths (*count of tenths*; e.g., responding 0.1 or 0.2 for $\frac{1}{2}$). Two other erroneous patterns corresponded to a decimal number using both components of the original fraction (*decimal reading*; e.g., responding 1.2 or 2.1 for $\frac{1}{2}$) or only one of them (*decimal reading* *with repetition*; e.g., responding 1.1 or 2.2 for $\frac{1}{2}$). A final pattern that we detected corresponded to a mixed number whose whole part and denominator are components plausibly taken from the initial fraction (*mixed number reading*; e.g., responding 1.5, corresponding to $1 + \frac{1}{2}$, for $\frac{1}{2}$).

Each error pattern could be split into a main pattern operating on the target fraction, and a variant corresponding to the same operation performed on the inverse fraction (*inversion variant*). Both variants are described with their formulas in Table 2. In some cases, applying these patterns would have yielded responses inapplicable to the 0–5 decimally graduated line that we used. As an example, the inversion error pattern predicted for fraction $\frac{6}{2}$ a response falling on 0.3333…, which was not a valid decimal graduation. Similarly, the relative error pattern predicted for fraction $\frac{6}{3}$ an answer of 10, which fell outside the 0–5 range and was thus also not possible. We ignored these in the rest of the analyses.

To confirm that the above patterns accurately capture the majority of children's responses, we plotted in the first panel of Fig. 5 the density of responses to each target fractions. Combined, our patterns could explain on average 68 % of total errors on a given fraction (min: 48 %; max: 89 %).

When unsure, participants could have selected random positions on the number line, possibly favoring integers. Therefore, it was important to demonstrate that each error pattern occurred more frequently than chance. To do this, we focused on responses that could be specifically and unambiguously explained by an error pattern. For each pattern, we excluded targets where the same response was predicted by another pattern, to avoid artificially inflating the number of responses of a pattern.

Table 2 shows the average proportion of responses, computed across fractions, that could be unambiguously explained by each pattern. The proportion varied widely between patterns: some accounted for around 20 % of responses on average (*correct*, *decimal reading*, *count of units*), while others accounted for only a few percent (e.g., *decimal reading with repetition*). However, it is important to note that chance level was generally very low. If subjects were responding purely at random, we would expect only about 2 % of responses to fall at any given position, as there were 51 potential positions on the line. For some positions, chance level could be even lower if participants were specifically avoiding them by default (or preferring other positions). Details of the statistical significance tests for each pattern are provided in supplementary materials. Overall, we found that all patterns were significantly represented for at least a subset of fraction targets. Detailed results for three specific response patterns—*relative to the line, counting the numerator*, and *decimal reading*—are presented in the last panels of Fig. 5.

*Roles of the numerator and denominator*

As described earlier, we observed pairs of patterns that exhibited similar behavior, with the only difference being that the roles of the numerator and the denominator were swapped between them. In other words, children could apply two *variants* of a pattern, depending on the roles they assigned to the two components. But why did this inversion occur? Our first hypothesis was that these errors stemmed from a stimulus-independent encoding level: some subjects did not understand the roles of the numerator and denominator, and regardless of the stimulus, had a certain probability of swapping their roles. However, this hypothesis could be rejected because the inversion errors varied depending on the target fraction. To demonstrate this, we measured the relative frequencies of paired variants and tested whether the observed bias toward either variant differed across fractions (Fig. 6). This analysis required both variants to predict a response that fell on a decimal graduation within the 0–5 line. Once again, we excluded targets for which one prediction was shared between several patterns, to ensure the robustness of the results. As a result, two patterns (*count of units* and *relative*) had no valid targets and could not be analyzed. For all patterns but one, we observed different biases across fractions (*Correct/Inverted*: $\chi^2(3) = 128.00, p < .001$; *Count of tenths*: $\chi^2(13) = 156.65, p < .001$; *Decimal with repetition*: $\chi^2(5) = 40.13, p < .001$; *Mixed number reading*: $\chi^2(1) = 79.15, p < .001$). As an example, children whose response was a division of one component by the other (*correct* or *inverted*) were more frequently applying the correct order on $\frac{6}{3}$ and $\frac{8}{4}$ than on $\frac{3}{6}$ and $\frac{4}{8}$. Here, children probably considered the

**Table 2**

**Terminology used to classify the main types of erroneous responses to a target fraction.** The formula encodes the value of the response, assuming the original fraction has the form a/b. The column labeled "% observed" gives the average frequency with which that response type was observed in our corpus. The p-value arises from a chi-square test relative to the average frequency with which the same response was observed when a different target was presented. L and R refer to the numerical value of the left and right boundaries of the line.

| Response type | Formula | Example for 3/6 | % observed | *p* value (FDR-corrected) |
|---|---|---|---|---|
| Correct response | a/b | 0.5 | 15.4 | .003 |
| Inverted | b/a | 2 | 5.6 | .12 |
| Correct relative to line | L + (R-L) x a/b | 2.5 | 5.9 | .1 (.04 for '2.5′) |
| Inverted | L + (R-L) x b/a | 10 | 4.8 | .18 (.18 for '2.5′) |
| Count units (numerator) | a | 3 | 21.1 | .007 |
| Inverted (denominator) | b | 6 | 12.7 | .02 |
| Count tenths (numerator) | a/10 | 0.3 | 6.9 | <.001 |
| Inverted (denominator) | b/10 | 0.6 | 2.2 | <.001 |
| Decimal | a + b/10 | 3.6 | 16.2 | .001 |
| Inverted | b + a/10 | 6.3 | 15.3 | <.001 |
| Decimal with repetition | a + a/10 | 3.3 | 1.3 | .006 |
| Inverted | b + b/10 | 6.6 | 0.8 | .01 |
| Mixed fraction | a + 1/b | 3.166… | 2.3 | 0.25 |
| Inverted | b + 1/a | 6.333… | 1.3 | 0.50 |

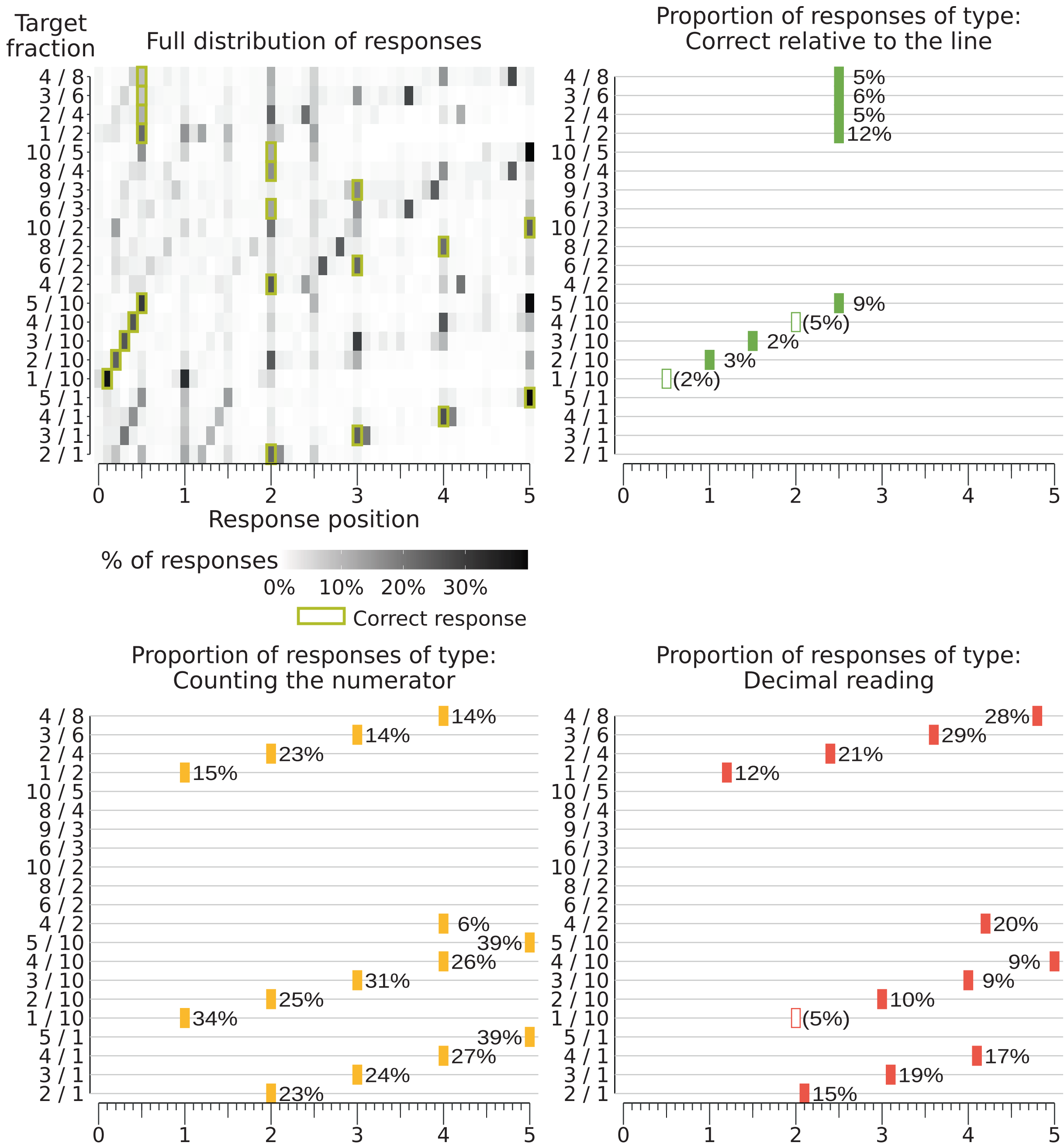


**Fig. 5. Distribution of all responses on the 0–5 number line.** On the left panel, darker tiles indicate a higher proportion of responses, while green boxes indicate the location of correct responses. Other panels show the proportion of responses respectively following the *relative* (top right), *count of units* (numerator; bottom left) and *decimal reading* (bottom right) error patterns. The numbers give the proportion of trials on which this response occurred for that target. Filled boxes and unparenthesized numbers indicate when that proportion was significantly higher than chance, as estimated by the probability of observing the same response when the target was different. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

difficulty of the division, swapping operands when the resulting operation was easier. The only pattern for which no difference was observed was the *Decimal* pattern ($\chi^2(5) = 5.90$, $p = .32$).

## *Discussion*

Our results show that French 6th graders are highly prone to errors, with a 78 % error rate, even on fractions as simple as $\frac{1}{2}$ or $\frac{2}{1}$,

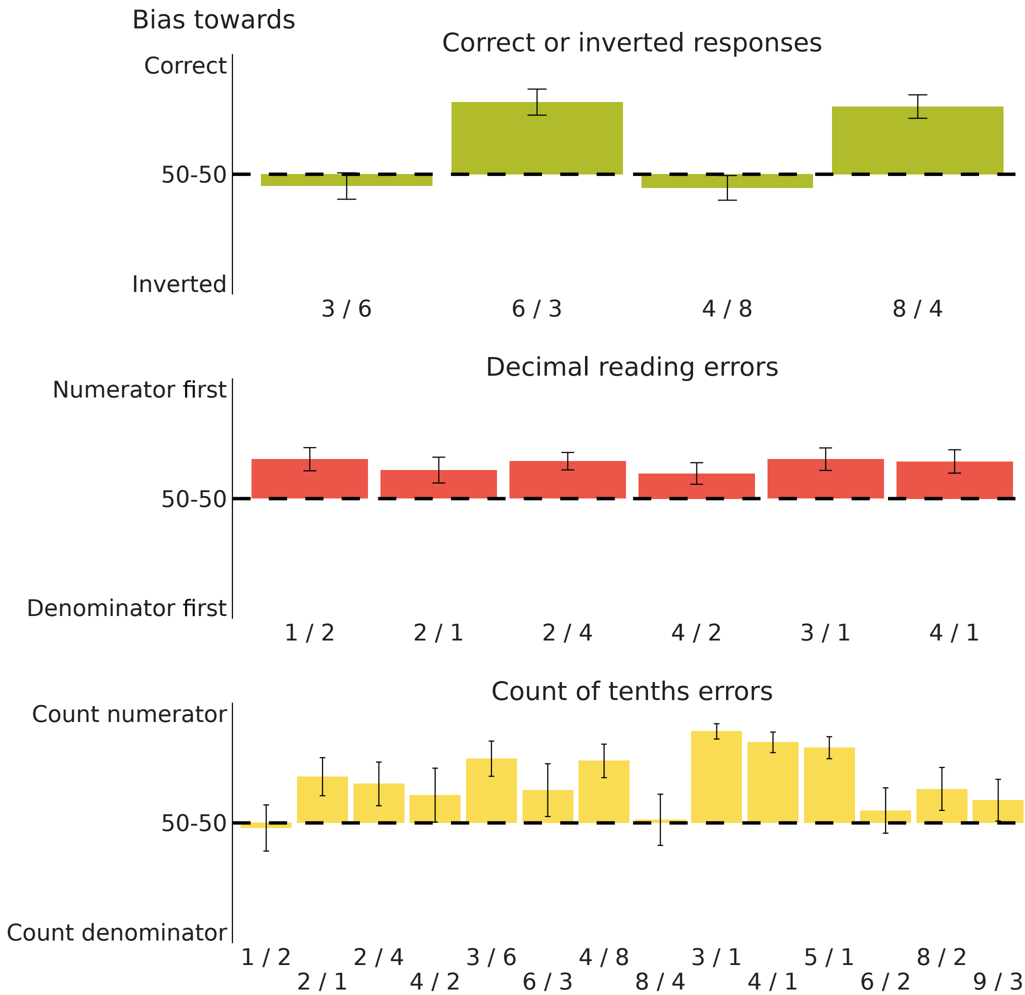


**Fig. 6. Probability of respecting the order of the numerator and denominator.** For each of three response types – *correct/inverted* (top panel), *decimal reading* (center panel) and *count of tenths* (bottom panel) – a bar indicates how frequently children followed the normal top-to-bottom ordering of numerator and denominator. A bar at the 50/50 level indicates random behavior, while a bar above that level indicates a trend towards a proper understanding that fractions must be read from top to bottom. Error bars indicate standard error.

despite the curriculum specifying that they should have studied fractions for the last two years. These errors are driven by a small number of response patterns (Bright et al., 1988; Hannula, 2003), and each pattern has two variants, depending on how the roles of the numerator and the denominator are interpreted. This results in similar distributions of responses between a target fraction and its inverse (Bright et al., 1988). Finally, those inversion errors do not occur randomly, but are more frequent for some target fractions than others. All these are basic observations that we now seek to interpret and understand.

Following Brown and colleagues (Brown & Burton, 1978; Brown & VanLehn, 1980), we sketched an algorithm (Fig. 7) describing how a proficient student would place a fraction on the line, which may explain error patterns as local deviations from this algorithm (or "bugs" (Van Lehn, 1990)). We propose that, for a given target, students may follow one of two main strategies: a *conversion strategy* and a *partition strategy*. In the conversion strategy, participants attempt to transform the target fraction into a number easier to manipulate, such as a decimal number that can be easily located on the graduated line. When performed correctly, this strategy corresponds to a division of the numerator by the denominator. In the partition strategy, the fraction is considered in its multiplicative meaning: $\frac{a}{b}$ means $a$ units of size $\frac{1}{b}$, or "$a$ $b$ths". Participants who understand this attempt, first, to identify the unit on the number line (it should be a segment of length 1); then to determine the size of a subsegment that divides this unit into b equal parts; and finally to find the location corresponding to $a$ times this amount starting from 0 (possibly by counting).

One key difference between the two strategies is the order in which the three elements of the problem are combined (the numerator, the denominator and the number line). In the conversion strategy, the numerator and the denominator are merged into a single number independently of the number line, before their response location is searched for. In the partition strategy, the denominator is first used to divide the line into subparts before taking into account the numerator. In Fig. 7, we propose a decomposition of each strategy into a series of intermediate goals. We expected the error patterns to correspond to incorrect executions of individual steps (Brown &

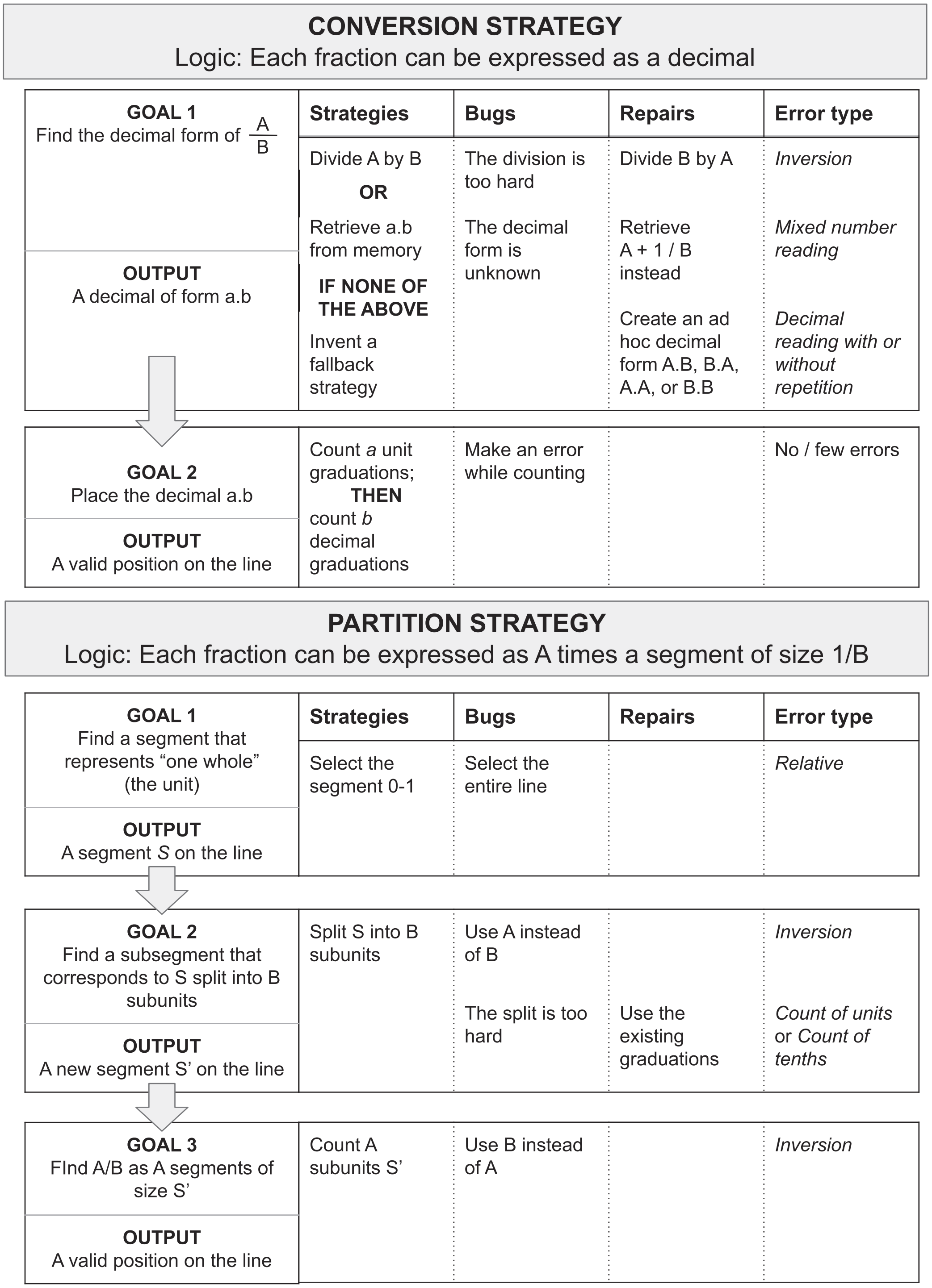


**Fig. 7.** Tentative two-strategy model of students' behavior.

VanLehn, 1980), which would still align with the expected input–output requirements.

*The conversion strategy and its bugs*

The number conversion strategy features two intermediate goals: converting the target fraction into a number easier to manipulate (an integer or a decimal), then placing it on the line. To convert a fraction into a decimal, an expert may simply rely on memory, especially for common fractions such as $\frac{1}{2} = 0.5$. However, during this stage of arithmetic fact retrieval, children may end up retrieving the result for another, related number (Ashcraft, 1992; Campbell, 1995), for instance the result of a mixed number whose name is close to the target (e.g., confusing "one half" and "one and a half"; this may explain *mixed number reading*). If the result is not known by heart, then the expert would divide the numerator by the denominator, which works for any fraction. However, students may find some divisions simpler than others, especially those that result in whole numbers. A student stuck in front of a seemingly impossible division (e.g. 3 divided by 6) might invert the numerator and the denominator to "repair" it (Brown & VanLehn, 1980) and have a calculation they can manage. This, too, would account for the *inversion* pattern.

What happens, however, if a child can neither retrieve the result from memory nor repair the target division? Since they are trying to achieve their main goal, i.e. find the decimal equivalent to the fraction, they may come up with an *ad hoc* fallback procedure by generating a decimal number using the component integers of the target fractions. Not knowing how to fill the *x.y* decimal frame they are searching for, they would fill it with whatever numbers are available in the stimulus, e.g., responding 1.2 to $\frac{1}{2}$. In doing so, they might use both the numerator or the denominator, in either order, thus explaining *decimal reading* errors, or a single unit of these numbers, thus accounting for *decimal reading with repetition*.

Once the decimal number is determined, the next step is to place it on the number line. The data suggest that this step is relatively error free, with only 13 % of errors occurring during the placement of decimal numbers presented as such. However, it is worth noting that the higher cognitive load in converting fractions to a decimal may have contributed to a slight increase in errors.

*The partition strategy and its bugs*

The partitive strategy features three subgoals to place a fraction of type $\frac{a}{b}$: identifying the unit (1), partitioning it into $b$ equal parts that serve as subunits (halves, thirds…), and then counting $a$ of those parts. According to this strategy, a fraction like $\frac{7}{5}$ is literally understood as 7 'fifths', without calculation, by determining the specific length corresponding to one fifth on the proposed number line.

An expert implementing this strategy would first identify the unit as a segment of length 1 on the proposed number line, here the 0–1 segment. Non-experts might mistakenly take the entire line as the unit, leading to a *relative* error pattern. Next, the expert would then divide the unit into equal parts based on the denominator. Here, errors can occur if students mistakenly use the numerator for partitioning (*inversion*) or rely on the existing decimal graduations, which mark a division of the unit in ten, resulting in a *count of tenths* error pattern. They might also use the unit graduations instead (*count of units*). Finally, the expert would count $a$ of parts of size $\frac{1}{b}$, obtained after the previous step. At this stage, errors can arise if students use the denominator for counting. The counting procedure itself should not pose any difficulty for 6th graders.

This model can help explain why children give similar answers on inverted fractions. As described above, some procedures may be particularly difficult to perform on a given fraction (e.g., dividing 3 by 6): our proposal is that these difficulties are bugs that are sometimes repaired, obviously inappropriately, by swapping the roles of the numerator and of the denominator (dividing 6 by 3, which results in 2). For instance, among the responses on $\frac{6}{3}$ that were either the correct one or its inverse, 83 % were correct; for $\frac{3}{6}$, this figure decreased to 44 %, arguably because the correct division was not round and thus, again, harder to perform. Our model can explain why *inversion* errors occur more frequently when the inverse fraction is round (the underlying division is simpler after swapping), and why this is not the case for decimal-reading errors: replacing the fraction bar by a decimal point is equally easy regardless of the specific integers in the original fractions.

Although this two-strategy model can explain all the patterns we observed, it remains speculative. Nonetheless, we think it provides valuable insights to explaining students' error patterns. In particular, this model applies Brown & VanLehn's repair theory (1980), originally developed for arithmetic procedures (Tatsuoka, 1984), to errors made on number-to-line tasks (Bright et al., 1988; Hannula, 2003). This suggests that procedural bugs (Brown & Burton, 1978) may underlie several observations from the earlier literature on number-to-line tasks. For example, Siegler and Thompson (2014) asked students to explain how they estimated the position of a fraction, and reported four main strategies: *numerical transformations* ("translating the fraction to a decimal or percentage"); *number line segmentation* ("imposing subjective landmarks on the number line"); *magnitude* strategies ("relying on the size of the fraction relative to the numerical range"); and *independent components* strategies ("[providing] estimates based solely on the numerator or denominator"). The first two strategies are identical to those proposed in our model. We propose that the *independent components* strategy is actually a "buggy" implementation of *number line segmentation* (the segmentation of the line is chosen irrespective of the denominator). This could explain why subjects who report using this strategy in Siegler and Thompson's study had a lower accuracy. Our model does not explicitly account for the *magnitude* strategy, likely because our task required the exact position of the fraction, making approximate size estimation less relevant.

## Study 2: The number line test in 8th and 10th grades

In Study 1, data from 6th graders allowed us to build a model of how students would place fractions on a number line. Study 2 served as a blind test for this model: could it account for new data from different age groups? Still in the context of France's national evaluations, we tested a representative sample of 20,000 8th and 10th graders. Although the French school curriculum assumes that fractions are mastered by 6th grade, the results from Study 1 suggested that students in later grades, too, would still make many errors on the task, albeit fewer. In addition, we expected their responses to exhibit the same error patterns predicted by our model and observed in Study 1.

*Material & methods*

*Procedure*

The procedure was identical to that of Study 1.

*Stimuli*

The stimuli were identical to that of Study 1.

*Participants*

In addition to our 6th graders sample, 5071 8th graders and 17,150 10th graders were tested at the beginning of the same school year, in September 2022. In the French school system, all students follow a common curriculum until 10th grade, after which they are split in three major pathways:

1. *Certificat d'Aptitude Professionnelle* (applied-10th grade), a two years program aimed at specific professions (e.g., cooks, hairdressers, carpenters).
2. *Seconde Professionnelle* (professional-10th grade), a three-year track that prepares students for the professional *Baccalauréat*, covering broader fields (e.g., administrative management).
3. *Seconde Générale et Technologique* (general-10th grade), the most general pathway, with a common curriculum covering literature, technology, and science.

The amount of time spent studying mathematics varies significantly in these pathways. Applied-10th graders are taught mathematics alongside physics and chemistry for 43.5 h, while professional-10th graders have 45 h specifically for mathematics. General-10th graders, on the other hand, receive 144 h of mathematics instruction. Most students (roughly two thirds), in particular good-

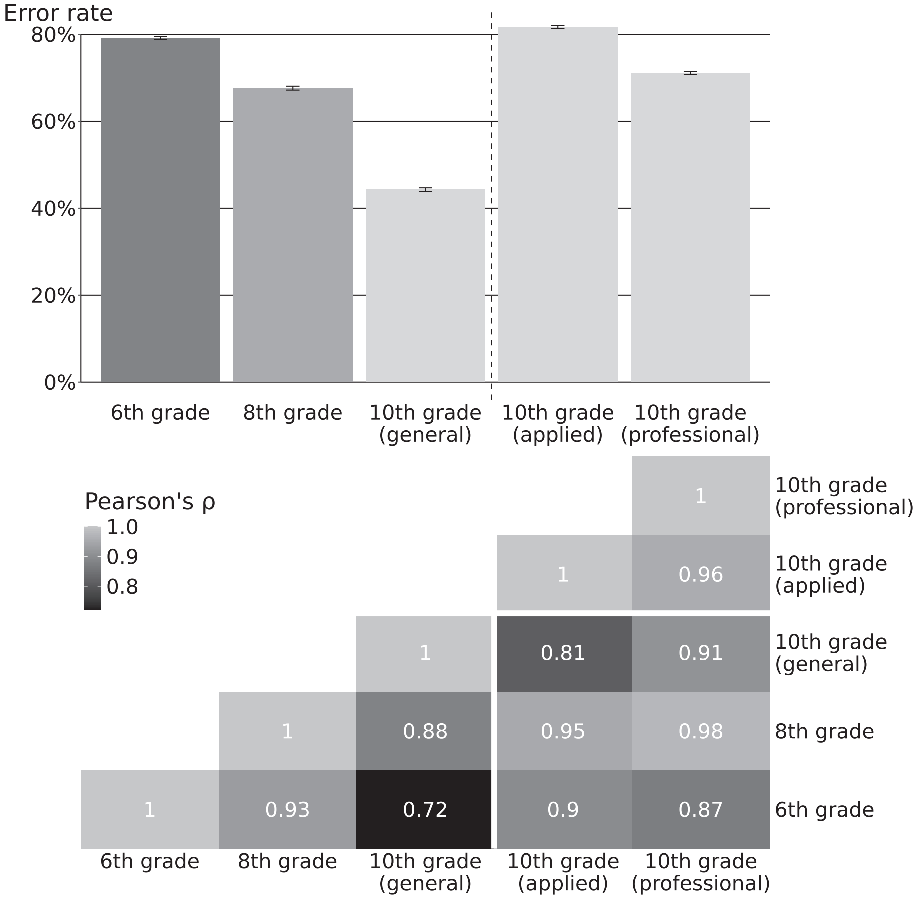


**Fig. 8. Results of experiment 2, where the same test was taken by 6th, 8th, and 10th graders.** Top: Proportion of errors on the number-to-line task by grade. The three bars at left are for students following the most common schooling trajectory (*6th-8th-general-10th grade*), while the two bars at right are for the two branching pathways in 10th grade (*applied* and *professional*). Error bars indicate standard errors. Bottom: correlation between response distributions those five groups.

performing ones, head into general-10th grade. We tested students in the first year of each curriculum. Each sample was selected to be representative of the tested grade and curriculum. Table 1 reports the mean age of participants as well as final samples for each grade as well as figures from each exclusion criterion. Again, post-hoc analyses revealed that the distribution of responses was highly correlated between groups with and without feedback (all Pearson's ρ between 0.97 and 0.998).

## Results

As in 6th grade, we first confirmed that students performed similarly, independently of the set of fractions they were assigned to. Once again, we found the distributions of responses to be extremely highly correlated across groups, with $R^2$s ranging from 0.97 in applied-10th grade to 0.99 in general-10th grade.

### Error rates in upper grades

Since 8th and 10th graders took the same test as 6th graders, we were able to directly compare their error rates, which are reported for each class in Fig. 8. We ran a linear model predicting error rate by grade level for the main curriculum (which covered 6th, 8th, and

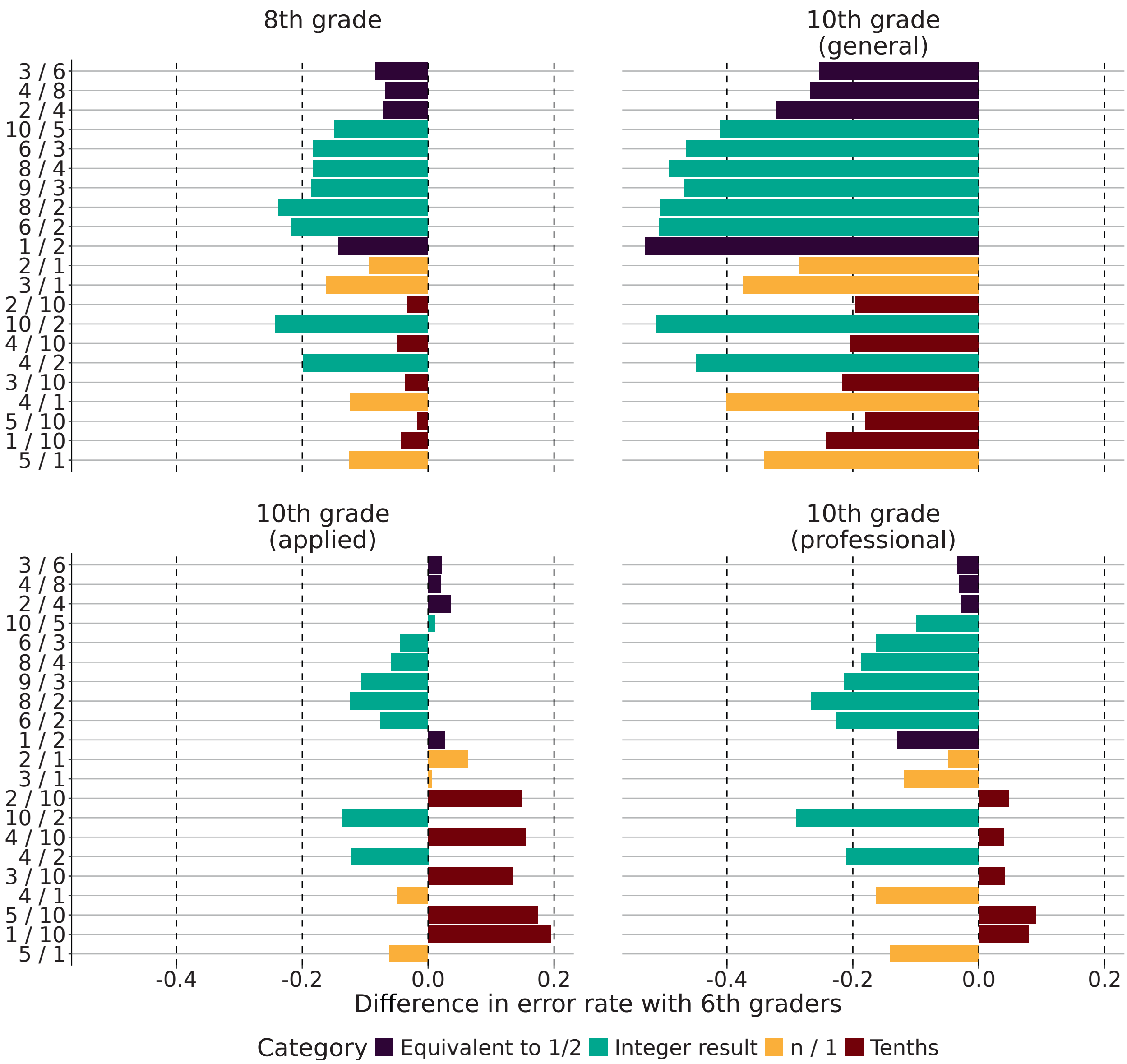


**Fig. 9. Differences in error rates for each target fraction with respect to 6th grade.** The panels give results for 8th grade (top left panel), general-10th grade (top right), applied-10th grade (bottom left) and professional-10th grade (bottom right). Colors indicate the type of fraction as defined in figure 2 (see legend at bottom).

general-10th grade; left side of the dashed line in Fig. 8). The model fitted the data significantly ($F(1, 16311) = 4240.0, p < .001, R^2 = .21$) and confirmed that students higher up in the curriculum made fewer errors ($\beta = -0.09$, $t(16311) = -65.12, p < .001$). In this linear model, we used grade level as a continuous predictor since we expected fraction experience to increase linearly after 6th grade, which was the last time students received formal fraction instruction. However, a model predicting error rate by grade, a categorical variable, provided a similar fit ($F(2, 16310) = 2204.0, p < .001, R^2 = .21$), with error rate decreasing both in 8th grade ($\beta = -0.12$, $t(16310) = -20.04, p < .001$) and in general-10th grade ($\beta = -0.34$, $t(16310) = -65.18, p < .001$). Across students in the same year of school, general-10th graders made fewer errors (44 %) than students in professional-10th grade (71 %; Welch's *t*-test across subjects: $t(11573) = -49.41, p < .001$), who made fewer errors than students in applied-10th grade (82 %; $t(9640) = -20.89, p < .001$). Indeed, applied-10th graders (82 % of errors) were close to the level of the average 6th grader (79 % errors), while students in professional-10th grade (71 % errors) were at a close level to that of students in 8th grade (68 % errors). This is perhaps not surprising since students in general-10th grade are usually selected among the best-performing students in earlier grades.

*Comparison by target*

We examined whether, across grades, students improved on all fraction items or only on specific ones. Fig. 9 reports the difference in error rates between each class and 6th graders, for each individual target. Following the order of grade levels, we first compared 6th graders to 8th graders with a two-way ANOVA on the error rate by grade and fraction type (similar to those used in Experiment 1: whole number result; $\frac{n}{1}$; $\frac{n}{10}$; and equivalent to ½) after computing the average error rate for each fraction, as we did in Study 1. As expected, there was a significant effect of grade ($F(1, 34) = 33.32, p < .001, \eta^2 = 0.50$) and type of fraction ($F(3, 34) = 9.31, p < .001, \eta^2 = 0.45$), but also a significant interaction between the two ($F(3, 34) = 2.98, p = .045, \eta^2 = 0.20$). This meant that 8th graders did not improve uniformly on all fractions. We then compared 8th graders and to the whole of 10th graders. There was, again, a significant effect of grade ($F(3, 68) = 58.77, p < .001, \eta^2 = 0.72$) and category ($F(3,68) = 20.73, p < .001, \eta^2 = 0.48$), but no interaction between the two ($F(9, 68) = 0.35, p = .95, \eta^2 = 0.05$). This meant that later grades only differed in their baseline performance with fractions and did not specifically improve on a given type of fraction.

*Comparison of error profiles*

Although older students made fewer errors, we expected their errors to be similar to those of 6th graders. We tested this by computing the correlation between the distribution of responses on individual targets, excluding the correct response, for each pair of grades. The average correlation over all targets is displayed in the bottom panel of Fig. 8. All pairs displayed a moderate (general-10th grade & 6th grade; $r = .72$) to almost-perfect (professional-10th grade & 8th grade; $r = .98$) pairwise correlation, suggesting that, as we predicted, error patterns were highly similar overall.

*Comparison of strategies*

Given the similar error distributions, we expected older students' errors to be driven by the same patterns as described in Experiment 1. Indeed, using our patterns, we could explain on average 62% (applied-10th grade) to 65% (8th grade) of the errors made

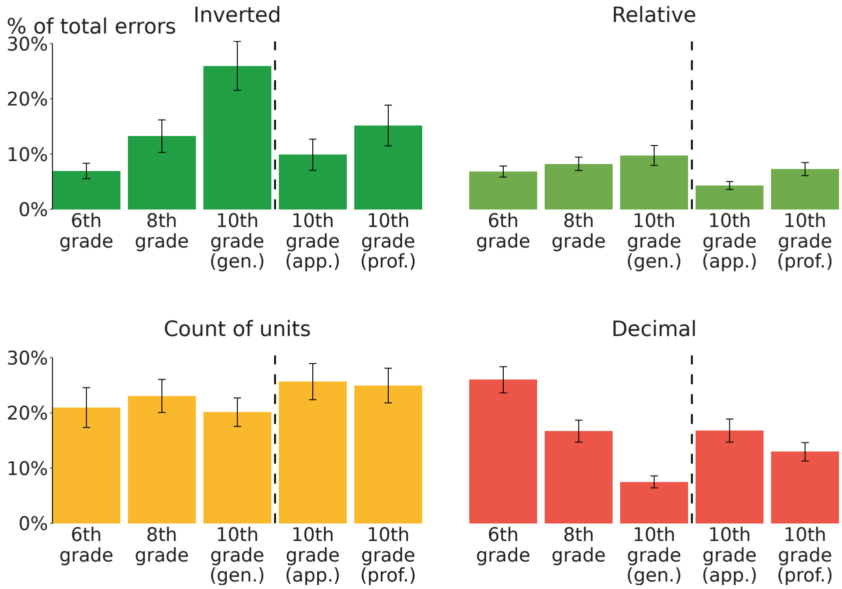


**Fig. 10. Comparison of pattern use across grades.** Each panel shows the relative proportions of errors explained by the four main error patterns in each grade: *inversion* (top left), *correct relative to the line* (top right), *count of units* (bottom left) and *decimal reading* (bottom right). Proportions were only computed over targets where the pattern was possible and the error could not be explained by another pattern. Colors match those used in figure 3. Error bars indicate standard errors.

on a given target, compared to 68% for the original sample of 6th graders. But we also expected differences across grades in the frequency of these error patterns. Although each pattern featured two variants, based on the role they assigned to the numerator and denominator of the target fraction, these variants arguably stemmed from very similar strategies. Therefore, we collapsed pairs of variants into a single pattern in the following analyses, with the exception of *correct* and *inverted* patterns which remained distinct. As before, we also ignored predictions which were common between two patterns.

We expected the main difference across grades to be in the proportion of errors explained by the different error patterns, which would change as students progressively adopted more correct strategies. For four patterns (*inversion*, *relative*, *count of units*, and *decimal*), we report in Fig. 10 the proportions of errors they explain in each grade, averaged over targets. To understand how the use of the patterns evolved with age, we ran for each pattern a linear regression, over targets, of the proportion of errors said pattern explained per grade level (6th-8th-10th grade). Although we had representative samples *within* each grade, general-10th graders, who were the best performers, were underrepresented among 10th-graders. To account for this, we weighted the 10th graders based on the distribution of students in 2022 (general-10th grade: 69.8 % of 10th graders; professional-10th grade: 22.7 %; applied-10th grade: 7.5 %) (Dauphin, Dieusaert, Juzdzewski, & Miconnet, 2022). Only *inversion* errors ($F(1, 48) = 19.85$, $p < .001$, $R^2 = .29$) and *decimal reading* errors ($F(1, 93) = 59.43$, $p < .001$, $R^2 = .39$) were significantly affected by age: as a proportion of their total errors, older pupils made more inversion errors ($\beta = 0.04$, $t(48) = 4.46$) but fewer decimal ones ($\beta = -0.04$, $t(48) = -7.70$). We also investigated the differences in pattern use between the three curricula in 10th grade. We compared these pairwise using a paired *t*-test over target fractions, independently for all patterns. Overall, results by education path matched those by age: general-10th graders, compared to professional-10th graders, made proportionally more *inversion* ($t(10) = 4.98$, $p < .001$) and *relative* errors ($t(13) = 3.15$, $p = .008$), but fewer *count of units* ($t(18) = -3.49$, $p = .006$) and *decimal reading* errors ($t(19) = -6.92$, $p < .001$). Professional-10th graders, compared to applied-10th graders, also made proportionally more *inversion* ($t(10) = 5.10$, $p < .001$) and *relative* errors ($t(13) = 4.89$, $p < .001$), but fewer *decimal reading* errors ($t(19) = -5.07$, $p < .001$). No difference was observed on *count of units* errors ($t(18) = -0.70$, $p = .49$). All p-values were FDR corrected across patterns and comparisons.

Overall, these results indicate a hierarchy of errors: some errors disappear across school years (e.g., *decimal reading*), while others become more and more frequent (e.g., *inversion*). This evolution is consistent within the framework of our model: errors come from buggy procedures that students refine over time and mathematical exposure. However, grade level is probably just an indirect estimate of how advanced each student is on this refinement scale. Our next analyses aimed to dissociate the influence of these two factors: grade and actual fraction performance.

*Individual variability: separating the effects of grade and performance*

Within each grade, there were important variations in performance level. We wondered whether the above trend (a progressive disappearance of decimal errors across grades, and the increased importance of inversion errors) could also be identified within a grade, based on individual accuracy. We separated students in five groups based on their number of correct responses on the five fractions: 0, 1, 2, 3 and 4 (students with perfect accuracy were left out since, by definition, their error patterns were undefined). For each pattern, we then tried to model its frequency of use across groups (the proportion of errors it could explain within each group) as a function of individual accuracy. Fig. 11 illustrates the tendencies that we tried to account for. As an example of what they mean, note

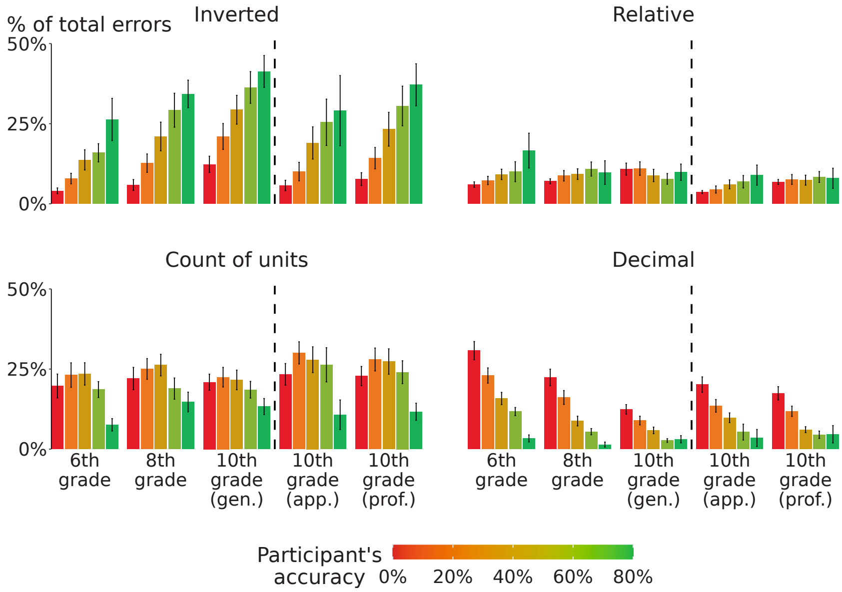


**Fig. 11. Characterization of inter-individual variability within each grade level.** Each panel shows the relative proportions of errors explained by one of the four main error patterns, now separated within each grade as a function of a child's total number of errors on the number-to-line task (accuracy, 5 levels shown in the color bar). Same format as in figure 10.

that in 6th graders with low individual accuracy (5 errors out of 5 targets), 30 % of errors were decimal errors, whereas in 6th graders with high individual accuracy (a single error out of five targets), only 5 % of errors were decimal errors – suggesting that, as performance improves, decimal errors decrease in a disproportionate manner. Again, these frequencies were computed over the fractions for which these errors were possible and unambiguously identifiable, which means that they cannot be directly compared across patterns.

For each pattern, we modeled the effect of individual accuracy on the pattern's frequency as a linear predictor across the 5 levels of students (1 to 5 errors). In all models, we controlled for average grade accuracy and grade level and their respective interactions with individual accuracy. Regressors were centered and reduced so that their effect sizes could be compared. Since there was an imbalance in the number of participants in the different groups (most students made 4 or 5 errors), we ran our models after averaging the frequencies over participants for each fraction, and then averaging again over fractions themselves –however, performing the same analysis over data averaged per participants yielded qualitatively similar results, albeit with a lower fit. The linear models provided significant fits for *inverted* errors ($F(5,19) = 172.60$, $p < .001$, $R^2 = .97$), *decimal reading* errors ($F(5, 19) = 96.87$, $p < .001$, $R^2 = .95$) and *relative* ($F(5, 19) = 23.03$, $p < .001$, $R^2 = .82$) errors. *Inverted* errors ($\beta = 0.10$, $t(19) = 26.17$, $p < .001$) and *relative* errors ($\beta = 0.10$, $t(19) = 5.60$, $p < .001$) were more frequent among total errors as students made fewer errors, while *decimal* errors were less frequent ($\beta = -0.06$, $t(19) = -18.54$, $p < .001$). The values for control variables are reported in Table 3. Although the fit was also significant for the *count of units* pattern, it was lower than for other patterns ($F(5, 19) = 3.68$, $p = .02$, $R^2 = .36$). In fact, Fig. 11 reveals that both students with low accuracy and students with high accuracy used it less frequently, and, accordingly, a model with both a linear and a quadratic regressor for individual accuracy confirmed this inverted U-shape pattern ($F(6, 18) = 31.05$, $p < .001$, $R^2 = .88$; *linear* effect: $\beta = -0.04$, $t(18) = -8.79$, $p < .001$; *quadratic* effect: $\beta = -0.05$, $t(18) = -9.27$, $p < .001$). Overall, individual accuracy was the best predictor of students' pattern use. Grade accuracy displayed a similarly oriented albeit smaller effect. The analysis confirmed and refined the hierarchy of errors we postulated earlier: *decimal reading* errors are dominant in students with the lowest level of understanding of fractions, followed by *count of units* errors and then by *relative* and *inversion* errors which occur predominantly among the highest-performing students.

## *Discussion*

Similar to the 6th graders tested in Experiment 1, 8th and 10th graders displayed a very low performance on the task. Even students from the best group of 10th graders made errors on 45 % of the trials, despite the French curriculum assuming full mastery of fractions magnitudes by 6th grade. In addition, 10th graders in applied and professional schools performed at a level comparable to respectively average 6th and 8th graders. Overall, 8th and 10th graders made the same types of errors as 6th graders, but their frequencies evolved, allowing us to sketch a putative developmental trajectory of fraction magnitude knowledge. *Decimal reading* errors, the most common error types in 6th grade, progressively disappeared, suggesting that they stemmed from a shallow misunderstanding of fractions. *Count of units* errors followed an inverted U-shaped curve, indicating that they were likely transition strategies (Rinne, Ye, & Jordan, 2017; Stafylidou & Vosniadou, 2004). Finally, *inversion* and *relative* errors became more frequent over time, suggesting that they arose from erroneous repairs in a more advanced strategy.

This ordering of error patterns could be observed not only across grades, but also within a grade, with better performing students making errors closer to that of the average student in a more advanced grade. Importantly, performance level was a stronger determiner of error types than grade. This large inter-individual variability in the understanding of fractions can be interpreted in several ways. One possibility is that some students need more years of education than others before attaining a given level of mastery. Another, non-necessarily incompatible one, is that some teachers or teaching methods may be more apt at conveying a proper understanding of fraction magnitude.

**Table 3**
**Individual and group predictors of pattern use across grades.** Each pattern was assessed independently using a linear model with the reported predictors. The square of individual accuracy was only included in the model for the *count of units pattern.* All regressors were normalized to allow comparing the coefficients.

| Pattern | Effects | | | | Interactions | |
|---|---|---|---|---|---|---|
| | Individual accuracy (linear) | Individual accuracy (squared) | Average accuracy of grade | Grade level (6–8-10) | Individual accuracy x grade level | Individual accuracy x average accuracy |
| Inversion | $\beta = 0.10$<br>$t(19) = 26.17$<br>$p < .001$ | – | $\beta = 0.04$<br>$t(19) = 8.99$<br>$p < .001$ | $\beta = 0.02$<br>$t(19) = 5.49$<br>$p < .001$ | $\beta = 0.007$<br>$t(19) = 1.61$<br>$p = .12$ | $\beta = 0.005$<br>$t(19) = 1.34$<br>$p = .20$ |
| Relative | $\beta = 0.01$<br>$t(19) = 5.60$<br>$p < .001$ | – | $\beta = 0.01$<br>$t(19) = 5.35$<br>$p < .001$ | $\beta = -0.01$<br>$t(19) = 8.99$<br>$p < .001$ | $\beta = -0.008$<br>$t(19) = -3.18$<br>$p = 0.005$ | $\beta = -0.009$<br>$t(19) = -3.83$<br>$p = .001$ |
| Decimal reading | $\beta = -0.06$<br>$t(19) = -18.54$<br>$p < .001$ | – | $\beta = -0.01$<br>$t(19) = -3.97$<br>$p < .001$ | $\beta = -0.03$<br>$t(19) = -7.22$<br>$p < .001$ | $\beta = 0.02$<br>$t(19) = 4.56$<br>$p < .001$ | $\beta = 0.007$<br>$t(19) = 2.05$<br>$p = 0.054$ |
| Count of units | $\beta = -0.04$<br>$t(18) = -8.79$<br>$p < .001$ | $\beta = -0.05$<br>$t(18) = -9.27$<br>$p < .001$ | $\beta = -0.01$<br>$t(18) = -3.39$<br>$p < .001$ | $\beta = 0.02$<br>$t(18) = 4.19$<br>$p < .001$ | $\beta = -\ 0.0006$<br>$t(18) = -0.18$<br>$p = .89$ | $\beta = 0.006$<br>$t(18) = 1.25$<br>$p = .23$ |

## General discussion

Many studies highlight that mastering fractions is a significant challenge for many students (Behr et al., 1983; Hannula, 2003; Larson, 1980; Ni & Zhou, 2005; Siegler & Pyke, 2013). Indeed, here we found that 6th, 8th and even 10th graders performed extremely poorly on a basic number-to-line task, which required simple fractions such as $\frac{1}{2}$ to be placed on a number line marked from 0 to 5. This was the case even though the French national curriculum requires fractions to be studied in 4th and 5th grade and to be consolidated by 6th grade; indeed, this curriculum, which is imposed to all public and most private French schools, requires the specific exercise of the number line to be introduced in 4th grade. This poor performance indicates either that the students had a very poor understanding of fractions, or that they were not able to leverage whatever understanding of fractions they possessed in order to solve the task at hand. Indeed, we cannot exclude that some aspects of the task implementation engendered difficulties that were not fully reflective of the students' genuine understanding—for example, there was some time pressure (a 10-second limit to answer) and the presence of graduations, factors which are both known to negatively affect children's precision on fraction estimation tasks (Fitzsimmons, Thompson, & Sidney, 2020; Siegler & Thompson, 2014). Nevertheless, the student's poor performance as well as the wild nature of their errors suggests that they had a limited understanding at least of the magnitude meaning of fractions, since the number-to-line task requires participants to think about magnitude—a key property that connects fractions to other numbers (Siegler et al., 2013). Had students been able to roughly estimate the magnitude of some of the target fractions, they should have responded with a location near the correct one, yet this was clearly not the case in our data. In particular, proper fractions, and to a greater extent, unit fractions should have been easily spotted as being smaller than 1, yet most responses to them were greater than this landmark, which shows that students failed to integrate fractions with integers on the basis of their magnitude. Even worse, the error patterns we identified suggest that many students applied procedures that treat the numerator and denominator as two independent numbers, sometimes in a fundamentally flawed way that can never produce a correct response. Therefore, it is safe to say that most students did not use magnitude to solve the task, which, in our view, characterizes a severe lack of fraction understanding. One could argue that students understood the concept of fractions and the magnitude estimation goal, but simply struggled with the computations required to attain this goal. However, note that our target fractions were always simple and required only knowing the very first entries in the multiplication tables ($2 \times 2 = 4$, $2 \times 3 = 6$, $2 \times 4 = 8$, $3 \times 3 = 9$ and $2 \times 5 = 10$), which should not pose any difficulty to normal 6th graders and above. More concerningly, many students (more than half in 6th grade!) provided wrong responses for fractions like 2/1 or 4/1, which required no actual computation.

Our work is novel in that it not only evaluates magnitude understanding in a large representative sample of students, but also attempts to classify children's errors by examining which exact position they chose on the line. From a detailed examination of students' responses, we identified seven main error patterns that explained together about two thirds of responses. Some patterns were linked to a conversion of the target fraction into another number (most often, a decimal naively formed by conjoining the two integers in the fraction); others with an incorrect partition of the line (e.g., counting the graduations, or dividing the whole line instead of the 0–1 segment). These two categories of pattern, conversion and partition, match children's (Siegler et al., 2011; Siegler & Thompson, 2014) but also adults' (Fitzsimmons et al., 2020) self-reports on their strategies for number-to-line estimation tasks.

Building on our observation of distinct error types, we built a tentative explanatory model. We proposed that all the patterns within a category can actually be explained as attempts to follow the proper procedure, yet with an erroneous step, as if a "bug" occurred within the student's behavioral program. In that regard, our model is a direct transposition of Brown & Van Lehn's (1980) repair theory to a domain beyond elementary arithmetic (Braithwaite & Siegler, 2023; Brown & Burton, 1978; Brown & VanLehn, 1980; Tatsuoka, 1984). One key interest of going beyond arithmetic is that participants no longer manipulate only abstract digit symbols (Braithwaite et al., 2017; Braithwaite & Siegler, 2023), but have to combine these manipulations with operations on the line, for instance splitting it in parts and counting them.

Our model suggests an integration of conceptual and procedural knowledge: while individual steps within a strategy are procedural in nature and may thus be poorly executed, the strategies themselves impose conceptual constraints on which procedures are acceptable. A student who uses the line partition strategy must know that a fraction has a multiplicative meaning that describes a number of parts, and thus has to–perhaps incorrectly–identify the parts and move through a certain number of them. However, the procedures available to them may not match the input they are provided with or the expected output. As an example, children may know a division procedure that can be used to convert a fraction into an integer when the dividend is a multiple of the divisor, but the conversion strategy may require more, as some fractions have a numerator smaller than the denominator. In that case, children may try to repair the known procedure by inventing a minimal adaptation of it (e.g., inverting the numerator and the denominator in the division process) (Brown & VanLehn, 1980; Van Lehn, 1990). If no procedure is available or appealing enough, children may instead invent one from scratch, still with the goal of matching the expected input/output requirements. This way of solving a task shares similarity with hacking, recently proposed as a model for human learning (Rule, Tenenbaum, & Piantadosi, 2020). Beyond the selection and creation of procedures, the choice of the strategy itself (numerical conversion versus line partition) most likely requires conceptual knowledge as well (Shrager & Siegler, 1998).

We expected both procedural and conceptual knowledge to evolve as students become more and more familiar with fractions; accordingly, we found an order in the evolution of errors both across grades but also within grades. *Decimal reading* errors, which were the most common error in 6th grade, were less frequent in older and better performing students. This is consistent with our view that these errors arise from a naive fallback procedure: as students become more familiar with fractions, they enrich their toolbox of procedures and no longer have to rely on invented ones. The errors that increase inform us on what these newly acquired procedures may be. *Count of units* errors, which are more frequent in 8th grade, correspond to a shift in strategy: students try to partition the line, but still fail to determine the proper size of the partition (which should be one over the denominator) and end up counting graduations. Although this procedure fails to account for the role of the denominator, it mimics a correct procedure when the denominator matches the graduations. As students start to acknowledge the role of the denominator, they should eventually move away from this procedure as well, explaining the inverted U-shaped evolution of its frequency. From there, students may reach a third level of understanding where errors now take into account the correct proportion expressed by the fraction. If students carry on with the line partition strategy, they may simply misunderstand the fraction as a proportion of the line, thus triggering *relative* errors (in which $\frac{1}{2}$, for instance, is placed in the middle of the 0–5 line, as *one half* of the line). But such errors remain rare, while *inversion* errors become dominant in general-10th grade and better performing students. This suggests that students are able to retrieve the correct ratio between the two components, most likely through a division, but do not take their order into account, perhaps for lack of attention to it, or because they do not manage to perform the division in the correct order and therefore fall back onto the inverse order. It is noteworthy that this later-developing understanding of multiplicative relationships is precisely what differentiates fractions from whole numbers and decimals. While whole numbers are typically interpreted through additive place value structures (e.g., $32 = 3 \times 10 + 2$), fractions rely on a fundamentally multiplicative relationship between numerator and denominator. In other words, decimal and count of units errors could be seen as an instance of the Whole Number Bias (Alibali & Sidney, 2015; Ni & Zhou, 2005), whereby children are overextending their knowledge from whole numbers. The task itself may facilitate the expression of this bias (Alibali & Sidney, 2015), since placing whole numbers on lines graduated in base 10 are part of the official curriculum up to 6th grade.

It is noteworthy that the three-level hierarchy was observed both within and across grades, largely driven by individual differences in performance. This suggests that, despite a standardized national curriculum, the French education system struggles to align students according to their grade level. In the future, it would be interesting to run the very same test in other countries such as Estonia, Portugal or Singapore, in order to see what can be achieved at each grade level by education systems renowned for their greater efficiency in math instruction.

Several questions remain unanswered. First, despite its consistency with our data, our model is only putative. Further testing is needed to probe the limits of its validity. In particular, our task only presented children with lines ranging from 0 to 5 and their answers were limited to graduations of step 0.1. This could have influenced children's strategy choices, e.g., when participants could not place $\frac{6}{3}$ at the position 6.3, which is the prediction made by the decimal error pattern. Testing children on different lines, new targets, or using less restrictive tasks, and describing how such variations affect children's behavior would be an important step in confirming our proposed model.

We also did not explore several alternatives to the strategies we described. As an example, students may apply another version of the line partition strategy in which they identify the unit, count it as many times as the numerator, and then split the result into as many parts as the denominator, thus executing the procedure we described in a different order. This procedure, which could not be separately identified in our sample, corresponds to the division meaning of fractions (e.g., faced with $\frac{5}{2}$, take 5 and split it in 2). Similarly, the line might be partitioned in not one but several steps: self-reports from the estimation task in Fitzsimmons et al. (2020) suggest that adults sometimes split the line first into halves and then further divide those halves into three to get sixths (which Fitzsimmons et al. labelled a *halves reference* strategy), instead of directly splitting the line into sixths (which would correspond to their *denominator* strategy). In the future, asking students to verbalize, as well as monitoring their eye movements, could facilitate the further distinction of sub-strategies.

More generally, although the two strategies described in our model conceptually align with the segmentation and transformation strategies described as the main strategies in the earlier literature involving number-to-line estimation tasks (Fitzsimmons et al., 2020; Siegler et al., 2011; Siegler & Thompson, 2014), their exact implementation differs. As an example, Fitzsimmons et al. (2020) reported three options for transformation strategies: equivalent simplification of a fraction (*simplifying*; e.g., converting $\frac{24}{84}$ as $\frac{6}{21}$); approximate simplification of a fraction (*rounding*; e.g., converting $\frac{54}{63}$ as $\frac{5}{6}$) and approximate conversion of a fraction into another number such as a decimal number (*translating*; e.g., converting $\frac{5}{8}$ as 0.6). Instead, our model offers a fourth option: conversion into an equivalent decimal number. This difference can be explained by the different natures of the two tasks: our model, based on an exact task, does not account for the approximation-based options. Additionally, since our analysis focused on errors and that equivalence knowledge is strongly related to accuracy on a number-to-line task (Fitzsimmons et al., 2020), it is also not surprising that we did not observe errors derived from a correct simplification of the target fraction followed by incorrect placement of the result. Nonetheless, this fourth option, if attested, would nicely complement the other three to form a 2x2 framework: approximate vs. equivalent (exact) conversion, and conversion into another fraction vs. a decimal number. Future work could aim at integrating these options, as well as other less frequent strategies, into a more general model able to cover both our number-to-line task and a more typical estimation task.

A broader model of fraction-to-line conversion should also incorporate factors related to the specific task design. Here, for instance, children's responses were snapped to the closest graduation, which could have led children to answer with the location right next to the one that they intended. Although the extremely low proportion of answers falling next to the correct answer suggests that this effect was minimal, taking such task-related errors into account might be important to model other future data. Children also had limited time to answer (10 s), which could have incentivized the use of shallow strategies that are faster to perform (e.g., decimal conversion) over correct ones (Fitzsimmons et al., 2020). Our task also featured graduations, which are known to affect children's performance when estimating fractions (Siegler & Thompson, 2014). It is thus possible that the students could have performed differently (better or worse) in an estimation task without graduations, and it would be interesting to check to what extent behavior on our number-to-line task is predictive of performance on a typical estimation task without graduations or integer landmarks. Replicating the observation of a poor performance on various tasks would help us identify whether students completely lack any proper understanding of fractions, or rather are distracted from using it because of the features of the task at hand.

Additionally, with only 5 fraction trials per participant, we could not assess whether strategy use was consistent within a given student. The design of a task even sometimes prevented consistency:, a child using the decimal pattern and answering 3.6 to target $\frac{3}{6}$ would not be able to answer 6.3 to $\frac{6}{3}$ as the line stopped at 5. Determining whether students are consistent is of major relevance because it could transform the present test into a tool for specific diagnosis and remediation. Further research is under way in our lab to evaluate whether, with more items, the number-line test may permit to precisely determine misconception(s) of fractions, at least as concerns their magnitude, and to provide learners with explanations specifically targeted to what they failed to understand.

Another limitation of our work is that we compared grades cross-sectionally and not longitudinally. Thus we could not measure how behavior on the task evolves across years. However, the consistency between the effects observed for grade, academic path, and individual accuracy suggests that our observations provide a first glimpse of the temporal evolution of fraction understanding. Again, in ongoing work, we are adopting a longitudinal perspective.

A final limitation of our study is that we only covered the range of 6th to 10th graders, while fractions are officially learned during 4th grade and keep being used even during adulthood. In light of the poor performance we observed in all grades, an outstanding question is to what extent adults finally manage to get a proper understanding of the magnitude of fractions.

## Conclusion

In our French sample, 6th, 8th and even 10th graders displayed a very poor ability to place simple fractions at their exact position on a graduated number line. Their poor performance reveals a small set of error patterns shared across individuals and grades. We propose that the patterns derive from two strategies with potentially flawed procedures, linking our model to the literature on bugs in children's arithmetic operations (Braithwaite & Siegler, 2023; Brown & VanLehn, 1980). We extend this idea beyond the range of arithmetic and, as a corollary, tackle the question of how procedural bugs integrate with conceptual knowledge. Our proposal also has the potential to bear significant implications for teaching. Identifying the errors made by students is a first step to correct them, and we think that our test could be transformed into a powerful diagnostic tool to capture a given student's understanding of fractions.

## CRediT authorship contribution statement

**Maxime Cauté:** Writing – review & editing, Writing – original draft, Visualization, Validation, Formal analysis, Conceptualization. **Cassandra Potier Watkins:** Writing – review & editing, Validation, Project administration, Formal analysis, Conceptualization. **Chenxi He:** Writing – review & editing, Validation, Formal analysis, Conceptualization. **Stanislas Dehaene:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Funding acquisition, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This work was supported by INSERM, CEA, Collège de France, Université Paris Saclay, and an ERC advanced grant “MathBrain” (grant number ERC-2022-ADG 101095866) to S.D; M.C. was supported by a doctoral grant from the École Normale Supérieure de Rennes. We gratefully acknowledge an intense collaboration with the DEPP (Département de l'évaluation, de la performance et de la prospective) of the French national education ministry, and particularly Thierry Rocher and Sandra Andreu, who developed several successive versions of the experimental software under our supervision and ran the test in more than a thousand classrooms, as well as Ronan Vourc'h and Stéphanie Mas, who helped us access and understand the data. We are also grateful to the editor and the two anonymous reviewers for their feedback.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2025.106373.

## Data availability

Preprint, supplementary material and code are available at https://osf.io/267s5/. Data is property of the DEPP and cannot be shared without their consent.

## 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

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44*(1–2), 75–106. https://doi.org/10.1016/0010-0277(92)90051-I

Behr, M. J., Lesh, R., Post, T. R., & Silver, E. A. (1983). In *Rational number concepts* (pp. 91–125). Academic Press.

Bhatia, P., Delem, M., Léone, J., Boisin, E., Cheylus, A., Gardes, M.-L., & Prado, J. (2020). The ratio processing system and its role in fraction understanding: Evidence from a match-to-sample task in children and adults with and without dyscalculia. *Quarterly Journal of Experimental Psychology, 73*(12), 2158–2176. https://doi.org/10.1177/1747021820940631

Braithwaite, D. W., Pyke, A. A., & Siegler, R. S. (2017). A computational model of fraction arithmetic. *Psychological Review, 124*(5), 603–625. https://doi.org/10.1037/rev0000072

Braithwaite, D. W., & Siegler, R. S. (2023). A unified model of arithmetic with whole numbers, fractions, and decimals. *Psychological Review*. https://doi.org/10.1037/rev0000440

Bright, G. W., Behr, M. J., Post, T. R., & Wachsmuth, I. (1988). Identifying fractions on number lines. *Journal for Research in Mathematics Education, 19*(3), 215–232. https://doi.org/10.5951/jresematheduc.19.3.0215

Brown, J. S., & Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science, 2*(2), 155–192. https://doi.org/10.1016/S0364-0213(78)80004-4

Brown, J. S., & VanLehn, K. (1980). Repair theory: A generative theory of bugs in procedural skills. *Cognitive Science, 4*(4), 379–426. https://doi.org/10.1016/S0364-0213(80)80010-3

Cai, L., Choi, K., Hansen, M., & Harrell, L. (2016). Item response theory. *Annual Review of Statistics and Its Application, 3*(1), 297–321. https://doi.org/10.1146/annurev-statistics-041715-033702

Campbell, J. I. D. (1995). Mechanisms of simple addition and multiplication: A modified network-interference theory and simulation. *Mathematical Cognition, 1*(2), 121–164.

Chan, W.-H., Leu, Y.-C., & Chen, C.-M. (2007). Exploring group-wise conceptual deficiencies of fractions for fifth and sixth graders in taiwan. *The Journal of Experimental Education, 76*(1), 26–57. https://doi.org/10.3200/JEXE.76.1.26-58

Dauphin, L., Dieusaert, P., Juzdzewski, L., & Miconnet, N. (2022). *Les élèves du second degré à la rentrée 2022* (22.39). DEPP.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*(3–6), 487–506. https://doi.org/10.1080/02643290244000239

Dotan, D., & Dehaene, S. (2013). How do we convert a number into a finger trajectory? *Cognition, 129*(3), 512–529. https://doi.org/10.1016/j.cognition.2013.07.007

Fazio, L. K., Kennedy, C. A., & Siegler, R. S. (2016). Improving children's knowledge of fraction magnitudes. *PLoS One1, 11*(10), Article e0165243. https://doi.org/10.1371/journal.pone.0165243

Fitzsimmons, C. J., Thompson, C. A., & Sidney, P. G. (2020). Do adults treat equivalent fractions equally? Adults' strategies and errors during fraction reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(11), 2049–2074. https://doi.org/10.1037/xlm0000839

Gabriel, F., Szucs, D., & Content, A. (2013). The mental representations of fractions: Adults' same–different judgments. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00385

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences, 4*(2), 59–65. https://doi.org/10.1016/S1364-6613(99)01424-2

Groupe de travail du Csen. (2019). *ÉvalAide, un dispositif scientifique de prévention des difficultés en lecture et en mathématiques en CP et en CE1*. Canopé. https://eduscol.education.fr/document/7736/download.

Hamdan, N., & Gunderson, E. A. (2017). The number line is a critical spatial-numerical representation: Evidence from a fraction intervention. *Developmental Psychology, 53*(3), 587–596. https://doi.org/10.1037/dev0000252

Hannula, M. S. (2003). *Locating Fraction on a Number Line. In International Group for the Psychology of Mathematics Education* (Vol. 3,, 17–24.

Ischebeck, A., Schocke, M., & Delazer, M. (2009). The processing and representation of fractions within the brain. *NeuroImage, 47*(1), 403–413. https://doi.org/10.1016/j.neuroimage.2009.03.041

Jacob, S. N., & Nieder, A. (2009). Notation-independent representation of fractions in the human parietal cortex. *Journal of Neuroscience, 29*(14), 4652–4657. https://doi.org/10.1523/JNEUROSCI.0651-09.2009

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), 1845–1864. https://doi.org/10.1037/a0016892

Kieren, T. (1976). On the mathematical, cognitive and instructional foundations of rational numbers. In *Number and measurement: Papers from a research workshop* (pp. 101–144). ERIC/SMEAC.

Larson, C. N. (1980). Locating proper fractions on number lines: Effect of length and equivalence. *School Science and Mathematics, 80*(5), 423–428. https://doi.org/10.1111/j.1949-8594.1980.tb09687.x

Lewis, M. R., Matthews, P. G., & Hubbard, E. M. (2016). Neurocognitive architectures and the nonsymbolic foundations of fractions understanding. In *Development of mathematical cognition* (pp. 141–164). Elsevier. https://doi.org/10.1016/B978-0-12-801871-2.00006-X.

Matthews, P. G., & Chesney, D. L. (2015). Fractions as percepts? Exploring cross-format distance effects for fractional magnitudes. *Cognitive Psychology, 78*, 28–56. https://doi.org/10.1016/j.cogpsych.2015.01.006

Matthews, P. G., Lewis, M. R., & Hubbard, E. M. (2016). Individual differences in nonsymbolic ratio processing predict symbolic math performance. *Psychological Science, 27*(2), 191–202. https://doi.org/10.1177/0956797615617799

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature, 215*(5109), 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

Nuraydin, S., Stricker, J., & Schneider, M. (2022). No transfer effect of a fraction number line game on fraction understanding or fraction arithmetic: A randomized controlled trial. *Journal of Experimental Child Psychology, 217*, Article 105353. https://doi.org/10.1016/j.jecp.2021.105353

Obersteiner, A., & Hofreiter, V. (2017). Do we have a sense for irrational numbers? *Journal of Numerical Cognition, 2*(3), 170–189. https://doi.org/10.5964/jnc.v2i3.43

Patel, P., & Varma, S. (2018). How the abstract becomes concrete: irrational numbers are understood relative to natural numbers and perfect squares. *Cognitive Science, 42*(5), 1642–1676. https://doi.org/10.1111/cogs.12619

Pinheiro-Chagas, P., Dotan, D., Piazza, M., & Dehaene, S. (2017). Finger tracking reveals the covert stages of mental arithmetic. *Open Mind, 1*(1), 30–41. https://doi.org/10.1162/OPMI_a_00003

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

Rule, J. S., Tenenbaum, J. B., & Piantadosi, S. T. (2020). The Child as Hacker. *Trends in Cognitive Sciences, 24*(11), 900–915. https://doi.org/10.1016/j.tics.2020.07.005

Shrager, J., & Siegler, R. S. (1998). SCADS: A model of Children's strategy choices and strategy discoveries. *Psychological Science, 9*(5), 405–410. https://doi.org/10.1111/1467-9280.00076

Siegler, R. S., Fazio, L. K., Bailey, D. H., & Zhou, X. (2013). Fractions: The new frontier for theories of numerical development. *Trends in Cognitive Sciences, 17*(1), 13–19. https://doi.org/10.1016/j.tics.2012.11.004

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14*(3), 237–250. https://doi.org/10.1111/1467-9280.02438

Siegler, R. S., & Pyke, A. A. (2013). Developmental and individual differences in understanding of fractions. *Developmental Psychology, 49*(10), 1994–2004. https://doi.org/10.1037/a0031200
Siegler, R. S., & Thompson, C. A. (2014). Numerical landmarks are useful—Except when they're not. *Journal of Experimental Child Psychology, 120*, 39–58. https://doi.org/10.1016/j.jecp.2013.11.014
Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology, 62*(4), 273–296. https://doi.org/10.1016/j.cogpsych.2011.03.001
Stafylidou, S., & Vosniadou, S. (2004). The development of students' understanding of the numerical value of fractions. *Learning and Instruction, 14*(5), 503–518. https://doi.org/10.1016/j.learninstruc.2004.06.015
Tatsuoka, K. K. (1984). *Analysis of Errors in Fraction Addition and Subtraction Problems. Final Report*. Kikumi K. https://eric.ed.gov/?id=ED257665.
Tian, J., Bartek, V., Rahman, M. Z., & Gunderson, E. A. (2021). Learning improper fractions with the number line and the area model. *Journal of Cognition and Development, 22*(2), 305–327. https://doi.org/10.1080/15248372.2021.1890603
Van Lehn, K. (1990). *Mind bugs. The origin of procedural misconceptions*. MIT Press.