

1 Neurofunctional plasticity in fraction learning: an fMRI training study

2  
3 Silke M. Wortha<sup>1,2</sup>, Johannes Bloechle<sup>3</sup>, Manuel Ninaus<sup>1,4</sup>, Kristian Kiili<sup>5</sup>, Antero Lindstedt<sup>6</sup>, Julia  
4 Bahnmueller<sup>1,7</sup>, Korbinian Moeller<sup>1,4,7,8</sup> & Elise Klein<sup>4,9</sup>

5 <sup>1</sup> LEAD Graduate School & Research Network, University of Tuebingen, Tuebingen, Germany

6 <sup>2</sup>Department of Neurology, Universitätsmedizin Greifswald, Greifswald, Germany

7 <sup>3</sup>Department of Psychiatry and Psychotherapy, University of Tuebingen, Tuebingen, Germany

8 <sup>4</sup>Leibniz-Institut für Wissensmedien, Tuebingen, Germany

9 <sup>5</sup>Faculty of Education and Culture, Tampere University, Tampere, Finland

10 <sup>6</sup> Faculty of Information Technology and Communication Sciences, Tampere University, Pori, Finland

11 <sup>7</sup> Centre for Mathematical Cognition, School of Science, Loughborough University, United Kingdom

12 <sup>8</sup> Individual Development and Adaptive Education Center, Frankfurt am Main, Germany

13  
14 <sup>9</sup> Université de Paris, LaPsyDÉ, CNRS, Sorbonne Paris Cité, Paris, France  
15  
16  
17  
18  
19

20 Corresponding author:

21 Silke M. Wortha  
22 Department of Neurology,  
23 Universitätsmedizin Greifswald,  
24 Ferdinand-Sauerbruch-Straße,  
25 17475 Greifswald  
26 Germany

27  
28 [silkemaria.wortha@med.uni-greifswald.de](mailto:silkemaria.wortha@med.uni-greifswald.de)  
29  
30  
31  
32  
33  
34  
35  
36  
37

## Abstract

38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57

**Background:** Fractions are known to be difficult for children and adults. Behavioral studies suggest that magnitude processing of fractions can be improved via number line estimation (NLE) trainings, but little is known about the neural correlates of fraction learning.

**Method:** To examine the neuro-cognitive foundations of fraction learning, behavioral performance and neural correlates were measured before and after a five-day NLE training.

**Results:** In all evaluation tasks behavioral performance increased after training. We observed a fronto-parietal network associated with number magnitude processing to be recruited in all tasks as indicated by a numerical distance effect. For symbolic fractions, the distance effect on intraparietal activation was only observed after training.

**Conclusion:** The absence of a distance effect of symbolic fractions before the training could indicate an initially less automatic access to their overall magnitude. NLE training facilitates processing of overall fraction magnitude as indicated by the distance effect in neural activation.

**Keywords:** fraction processing, number line estimation training, flow experience, numerical distance effect, fMRI

Word count: 150

## Introduction

58  
59 Over the last decade, research has shown repeatedly that understanding fractions is a  
60 crucial predictor of future achievement in higher mathematics [1–3]. However, despite intense  
61 research efforts in this area, children’s poor performance when it comes to handling and  
62 understanding fractions hardly changed [4–7]. For example, in 2015 Lortie-Forgues and  
63 colleagues found that only 27% of 8<sup>th</sup> graders in the United States were able to choose correctly  
64 the number closest to the result of a fraction addition problem out of four given solution probes  
65 [8]. A similar result was already reported in 1978 by the National Assessment of Educational  
66 Progress [9], when only 24% of the 8<sup>th</sup> graders chose the correct answer to the same question.  
67 Importantly, these difficulties in understanding fractions may be persisting regardless of  
68 educational efforts because - unlike natural numbers - fraction magnitude processing seems to  
69 be more difficult due to its bipartite structure reflecting the relative relation of numerator and  
70 denominator [10]. According to the integrated theory of numerical development, magnitude  
71 information is the crucial basis for understanding numbers. Moreover, the understanding that  
72 all real numbers (e.g., natural numbers, integers, rational numbers) can be represented on a  
73 number line is a key assumption for numerical learning. Therefore, promoting fraction  
74 magnitude understanding seems crucial for fostering fraction understanding more generally  
75 [11–13]. Thus, interventions with the aim to improve fraction understanding and therefore  
76 conceptual knowledge of fractions should focus on fostering mastery in processing and  
77 representing fraction magnitude. In the context of (fraction) magnitude understanding number  
78 lines have been shown to be a beneficial instructional tool [14].

79 Against this background, we aimed at understanding the neuro-cognitive foundations  
80 underlying successful fraction learning and their plasticity. For this purpose, participants had to  
81 complete a number line estimation training and a flow questionnaire on five consecutive days.  
82 In the following we will highlight the most important research results from the research areas  
83 which are relevant for our study. First, we will give a brief overview about the relevance of

84 number line estimation training for fraction magnitude understanding. Second, we will  
85 introduce flow as a state which is beneficial for learning and especially for fraction learning.  
86 Third, we will summarize present knowledge about the neural correlates of fraction processing  
87 and highlight the importance of our study in this context. Finally, we will introduce the aim of  
88 the current study including our specific hypothesis.

### 89 Number line estimation as predictor for fraction magnitude learning

90 The mental number line is a metaphor for the nature of the number magnitude  
91 representation whereby numbers are represented spatially with their magnitude increasing from  
92 left to right (at least in Western cultures [15]). In numerical cognition research, number line  
93 estimation (henceforth NLE) is used repeatedly to assess number magnitude understanding –  
94 especially in children ([16–18], but see [19] for additional aspects). In the NLE task,  
95 participants have to indicate the spatial position of a target number on a given number line for  
96 which only start- and endpoint are specified [18].

97 As magnitude is the semantic core for any type of number, the task can not only be used  
98 to assess, but also to train fraction magnitude understanding [12]. For instance, Hamdan &  
99 Gunderson [20] conducted a training study with three conditions for fraction learning (i.e.,  
100 number line estimation training, area model training, and a non-numerical control). They  
101 observed that although children in both the NLE training and the area model training improved  
102 in the respective tasks, only children completing the NLE training showed transfer effects to a  
103 not trained magnitude comparison task with fractions.

104 Moreover, Barbieri and colleagues [21] used a number line-based intervention to  
105 improve fraction understanding in children with poor conceptual knowledge of fractions and  
106 compared the number line intervention group to a standard mathematics curriculum group. The  
107 number line intervention group showed significantly larger learning gains than the control  
108 group. Finally, computerized and game-based versions of the NLE task were used successfully

109 to assess and improve children's fraction understanding [22–24]. Taken together this  
110 substantiates that number lines are a powerful instructional tool and the NLE task can be applied  
111 successfully to foster fraction (magnitude) understanding. However, successful fraction  
112 magnitude learning might not only depend on improving conceptual knowledge of fractions,  
113 but also on a more fundamental ability to be able to reach a beneficial cognitive state for  
114 learning which is known as flow experience.

### 115 *Flow experience as an indicator for optimal learning*

116 Learning is not a pure cognitive process but is affected by motivation and emotions [25].  
117 A beneficial factor for learning that is considered specifically in computerized approaches on  
118 learning is the *flow experience* of the learner [26]. Flow was first coined by Csikszentmihalyi  
119 [27] and can be described as a positive emotional state [28,29] and as a holistic approach to  
120 motivation [30]. In particular, flow is characterized by a combination of factors such as  
121 increased concentration, reduced self-consciousness, sense of control, that are experienced as  
122 intrinsically rewarding [29]. Flow is usually reached when task demands meet personal skills  
123 or resources in a balanced way. Thus, when the skills of the learner are too low for the demands  
124 of a given task – for instance at the beginning of a training – flow experience is rather low. The  
125 same is true when the skills of the learner are too high for a given task which leads to boredom  
126 and reduced flow experience. Therefore, flow experience seems to be an optimal state for  
127 intrinsically motivated learning, which helps focus on the given task and can lead to improved  
128 performance [31]. This is further supported by studies on the neural correlates of flow  
129 experience. These studies could show that flow experience is associated with increased  
130 activation in areas of the multiple demand system such as inferior frontal gyrus, putamen and  
131 anterior insula and decreased activation in areas typically associated with the default mode  
132 network such as amygdala, medial prefrontal cortex and posterior cingulate cortex [32,33].

133 Flow experience was specifically, but not solely considered in different computer-based  
134 learning settings. For instance, in game-based learning (for a review see [34]), collaborative  
135 learning of problem-solving in virtual environments [35], hypermedia learning [36], e-learning  
136 [37], but also in creative processes like music learning [38,39]. As such, flow experience should  
137 also be beneficial for fostering fraction magnitude understanding using a computerized NLE  
138 training. However, successful fraction magnitude learning might not only depend on improving  
139 conceptual knowledge of fractions and on the learners' flow experience, but also on the  
140 successful interplay of certain neural correlates underlying the neuro-cognitive foundations of  
141 fraction learning. Therefore, in the following the current state of research on neural correlates  
142 of fractional learning is briefly outlined.

143

#### 144 Neural correlates of fraction and proportion magnitude processing

145 Despite above described established relevance of fraction knowledge and longstanding  
146 research in educational sciences and psychology, little is still known about the neural  
147 mechanisms underlying the processing of proportions and fractions in general and the neural  
148 correlates of fraction learning in particular. To date, there are only few neuroimaging studies  
149 investigating the neural correlates of processing proportions [40–42] and fractions [43–46] in  
150 adults.

151 One important aspect across most studies is that the numerical distance effect was used  
152 as an indicator of automatic processing of overall fraction magnitude. The numerical distance  
153 effect [47] describes the finding that two numbers are compared faster and more accurately the  
154 larger the numerical distance between them (i.e., the farther apart they are on the mental number  
155 line, e.g. 3 and 7 is easier to compare than 3 and 4). For instance, for fraction magnitude  
156 comparison, Ischebeck et al. [45] observed that neural activation within the right IPS was  
157 modulated by the overall numerical distance between the to-be-compared fractions (e.g.,

158 numerical distance between  $\frac{2}{4}$  and  $\frac{3}{7}$ ), but not by the numerical distance between numerators  
159 or denominators (i.e., numerical distance between 2 and 3 for numerators and numerical  
160 distance between 4 and 7 for denominators when comparing  $\frac{2}{4}$  and  $\frac{2}{7}$ ). Moreover, Mock and  
161 colleagues [41] observed a joint neural correlate of specific occipito-parietal activation  
162 including the right intraparietal sulcus (IPS) for the processing of different notations of  
163 proportions including not only fractions, but also pie charts, dot patterns and decimals.

164 Finally, Klabunde et al. [48] conducted a first fMRI training study on proportions in  
165 participants with fragile X syndrome and a control group with intellectual disabilities.  
166 Participants were trained for two days in 10 min sessions until they were able to have over 80%  
167 accuracy on matching fractions to pie charts and pie charts to decimals. Neurofunctional  
168 changes from before to after the training indicated significantly increased brain activation in  
169 the left inferior parietal lobule, left postcentral gyrus, and left insula for both groups. However,  
170 the mechanism of interest in this study was not the distance effect but to investigate neural  
171 correlates of stimulus equivalence relations.

172 In summary, previous studies indicate that the distance effect for overall magnitude of  
173 the to-be-compared proportions/fractions seems a good measure reflecting automatic  
174 processing of overall fraction magnitude. As such, the numerical distance effect as a hallmark  
175 effect for magnitude processing indicated that the (right) IPS seems to play a crucial role in the  
176 processing of proportion and fraction magnitude independent of the actual task, which is in line  
177 with its involvement in the processing of natural number magnitude [49]. As such, one might  
178 argue that the presence of a numerical distance effect seems to indicate automatic processing  
179 of overall fraction magnitude in proficient fraction processing (see [50] for a similar argument  
180 on the relation between distance effect and arithmetic performance), while absence of a  
181 numerical distance effect might indicate that the fractions' magnitude is not automatically  
182 accessed.

183

184 The present study

185           Against this background, we aimed at investigating neuro-functional correlates and their  
186 plasticity associated with an NLE training of fraction magnitude understanding. In particular,  
187 we evaluated changes in fraction magnitude processing on the neural level as reflected by the  
188 numerical distance effect for overall fraction magnitude. To the best of our knowledge, this is  
189 the first study investigating the neural correlates of fraction learning through a NLE training in  
190 healthy adult participants. We assessed neural activation associated with fraction magnitude  
191 processing using fMRI before and after a five-day consecutive NLE training on fractions.  
192 Additionally, we assessed participants' flow experience in each training session.

193           Similar to previous studies applying NLE training to children, we expected the training  
194 to improve participants' conceptual knowledge of fraction magnitude on a behavioral level  
195 [20,21]. In particular, we expected significant improvements in NLE performance over the five-  
196 day training sessions. Additionally, we expected participants flow experience to be associated  
197 with NLE training improvements over the five-day training. Finally, on the neural level, we  
198 expected significant changes of IPS activation associated with the numerical distance effect  
199 from pre- to post fMRI session indicating more automatic processing of fractions' overall  
200 magnitude after the NLE training. This should become evident by a more pronounced numerical  
201 distance effect in behavioral measures but also neural activation in IPS after the training.

202

203

## Methods

### 204 **Participants**

205           48 right-handed adult participants ( $M = 23.73$ ,  $SD = 3.65$ , female = 32) took part in this  
206 fMRI study. All participants were German native speakers with normal or corrected to normal  
207 vision and reported no history of psychiatric or neurological disorders or drug abuse. The study  
208 was approved by the local Ethics Committee of the Medical Faculty of the University of  
209 Tübingen. Participants gave written informed consent and received monetary compensation for  
210 their participation.



## 211 **Study Design**

212           The study was designed as a pre-post fMRI training study with five consecutive days of  
213 training between the fMRI measurements. On the first day before the training and on the last  
214 day after the training, participants had to complete two different magnitude comparison tasks  
215 (i.e., comparison of symbolic fractions and comparison of line proportions, respectively) and a  
216 fraction-line proportion matching task (i.e., indicating whether the magnitude reflected by a  
217 fraction matched that of a line proportion or not) while their brain activation was measured  
218 using fMRI (see Figure 1). In addition to these computerized tasks, participants also completed  
219 a paper-pencil-based NLE task prior to entering the scanner for pre- and post-test measurement  
220 (for more details see below). Due to technical problems, behavioral data of fMRI measurements  
221 could only be obtained from 24 right-handed adult participants ( $M = 22.50$ ,  $SD = 3.76$ , female  
222 = 16). Imaging results did not differ substantially between the participant group with and  
223 without behavioral data (i.e., no suprathreshold clusters at an uncorrected  $p < .001$  and cluster  
224 size of 10). Therefore, imaging results as well as all behavioral data obtained outside the scanner  
225 (i.e., training data, flow experience and paper-pencil-based number line estimation task) will be  
226 reported for all 48 participants, while behavioral fMRI results are reported for the respective 24  
227 participants only.

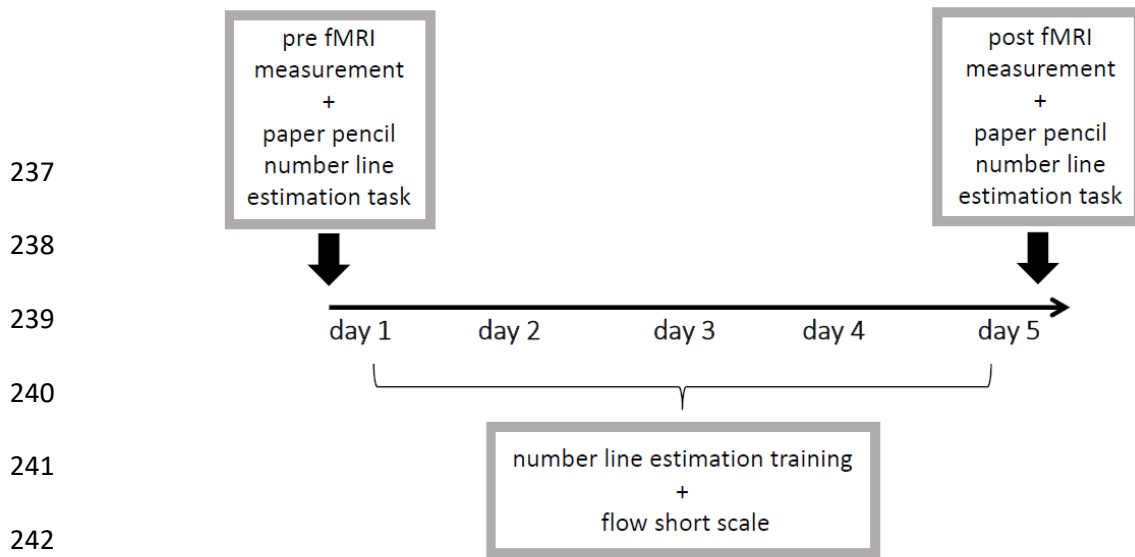
228           The training consisted of a fraction number line estimation task. It took place outside  
229 the scanner and each training session lasted around 15-20 minutes depending on participants  
230 individual performance. After each training session participants completed a brief questionnaire  
231 on mental flow to evaluate possible changes in flow experience over the training period (for  
232 more details on the flow questionnaire see below).

233

234

235

236



243  
 244 **Fig. 1: Study design.** Before and after a five-day number line estimation training fMRI measurements were  
 245 conducted to evaluate neurofunctional changes in brain activation through the training. After each training session  
 246 flow was assessed using the flow short scale [51]. Prior to entering the scanner for the pre and post measurement  
 247 a paper-pencil number line estimation task was administered.  
 248

### 249 **Item sets and stimuli**

250 In this study, two different item sets (one to be trained and one not to be trained) were  
 251 used. The order of sets was counterbalanced across participants so that whichever set a  
 252 participant did not train on served as the untrained set at the pre- and post-test. However, each  
 253 participant was tested during the fMRI sessions on both item sets. Each item set consisted of  
 254 48 stimuli with items always presented in randomized order. To ensure comparable difficulty  
 255 of item sets, they were matched on overall problem size and numerical distance between  
 256 fraction pairs used. All fractions used in the stimuli sets were proper fractions with nominators  
 257 and denominators ranging from 1 to 30. To evaluate neurofunctional changes in fraction  
 258 magnitude processing, pre-post-test evaluation tasks performed in the fMRI scanner included  
 259 both trained and untrained items. Items of the two item sets can be found in the Appendix.

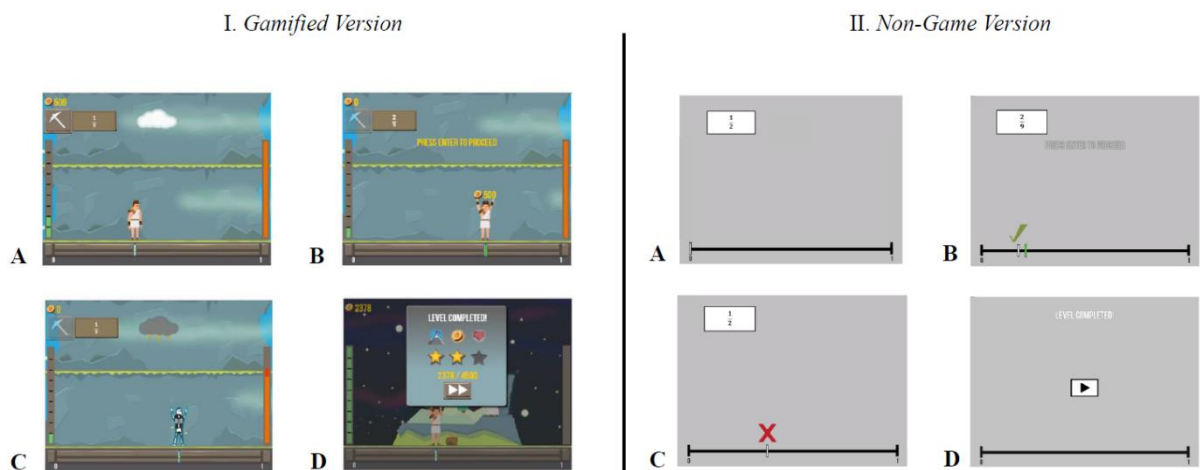
### 260 **Training Procedure**

261 For the training, we used two variations of a fraction number line estimation (NLE)  
 262 training with feedback, which differed in appearance and framing, but not in numerical content  
 263 and task. In particular, half of the participants were trained using an NLE task set within a  
 264 *gamified* environment (see Figure 2), while the other half were trained in the same NLE task in  
 265 a *non-game* environment (see Figure 2). For the gamified NLE task we used the *Semideus*

266 research engine, which was already applied successfully in previous studies for assessing and  
 267 training fraction knowledge [23,24,52–54].

268 In both versions of the NLE task, participants had to indicate the correct position of a  
 269 given fraction on a number line ranging from 0 to 1 by maneuvering an avatar in the gamified  
 270 version and moving a white slider along the number line using the left and right arrow keys of  
 271 a computer keyboard in the non-game version. After reaching the estimated correct position,  
 272 participants had to press the space bar to select that position. Participants were instructed to  
 273 indicate the right position as quickly and accurately as possible within a time limit of 10  
 274 seconds. They received positive feedback (i.e., cheering avatar plus coins awarded in the  
 275 gamified vs. green checkmark in the non-game version) when their answer fell within a range  
 276  $\pm 5\%$  around the correct position. In case they failed to answer or did not answer accurately  
 277 enough, negative feedback was given (i.e., avatar struck by lightning plus loss of virtual energy  
 278 in the gamified vs. red cross shown in the non-game version) and participants had to repeat the  
 279 item until it was correctly solved within the  $\pm 5\%$  range. At each training session, participants  
 280 worked through 96 items in 12 runs containing 8 items each. Each item from the trained  
 281 stimulus set was encountered twice within a training session. Items were presented in  
 282 randomized order and were identical in both versions of the NLE training.

283



284

285 **Fig. 2: Examples of different stages of the gamified (I.) and non-game version (II.) of the NLE training.**  
286 I. *Gamified version*: (A) Beginning of a new trial. A fraction is shown in the left upper corner and the avatar  
287 Semideus has to be moved on a number line between 0 and 1 towards the anticipated position on the number line  
288 of the shown fraction. (B) The result of a successful trial. If the position of the fraction is estimated correctly  
289 (tolerated range:  $\pm 5\%$ ) Semideus is rewarded with coins. (C) The result of a failed trial. In case the position of the  
290 fraction is not estimated correctly Semideus is struck by lightning and the participant needs to try again. (D)  
291 Completed level with feedback. Mountain: For completing the level, they earned one star; Coin: for collecting at  
292 least 2000 points they earned a second star; Heart: for maintaining at least 80% of the life points, they earned a  
293 third star. II. *Non-game version*: (A) Beginning of a new trial. A fraction is shown in the left upper corner and the  
294 participant has to move the white slider on a number line between 0 and 1 towards the anticipated position on the  
295 number line of the shown fraction. (B) The result of a successful trial. In case the position of the fraction is  
296 estimated correctly (tolerated range:  $\pm 5\%$ ) a green check mark appears. (C) The result of a failed trial. If the  
297 position of the fraction is not estimated correctly a red cross appears and the participant must try again. (D)  
298 Completed level.  
299

### 300 **Flow Short Scale**

301 Flow was assessed using the German version of the flow short scale [51]. This  
302 questionnaire consists of 16 items. Thirteen items are associated with the flow scale (7-point  
303 Likert-scale ranging from 1 = totally disagree to 7 = totally agree), which has a three-  
304 dimensional structure: The first dimension assessed by the scale is *fluency of performance* (6  
305 items, i.e., “My thoughts or activities run fluently and smoothly”,  $\alpha = .92$ ). The second  
306 dimension is *absorption by activity* (4 items, i.e., “I’m completely focused on what I’m doing  
307 right now”,  $\alpha = .80$ ). Finally, the third dimension is *perceived importance* or concern (3 items,  
308 i.e., “I’m worried about failure”,  $\alpha = .90$ ).

309 Additionally, the questionnaire includes 3 items of a demand scale (10-point Likert-  
310 scale ranging from 1 = easy to 10 = difficult), which aims at assessing how demanding the  
311 current task was for the participant (i.e., “For me personally, the current requirements are....”).  
312 To monitor flow experience over the course of the training, participants had to complete the  
313 flow short scale following each completed training session resulting in five completed flow  
314 short scales per participant.

### 315 **Paper pencil number line estimation task outside the fMRI scanner**

316 In addition to the computerized in-scanner tasks (see below) participants had to  
317 complete a paper pencil version of a number line estimation task prior to both fMRI  
318 measurements (pre and post training), respectively. The number line ranged from 0 to 1 with

319 only the endpoints specified and participants had to indicate the spatial position on the number  
320 line for all items from both item sets (i.e., the trained and the untrained set, thus 96 items in  
321 total prior to and after the training). This task allowed to evaluate potential improvements in  
322 spatial localization of fractions on the number line through the training.

### 323 **Tasks performed inside the fMRI scanner**

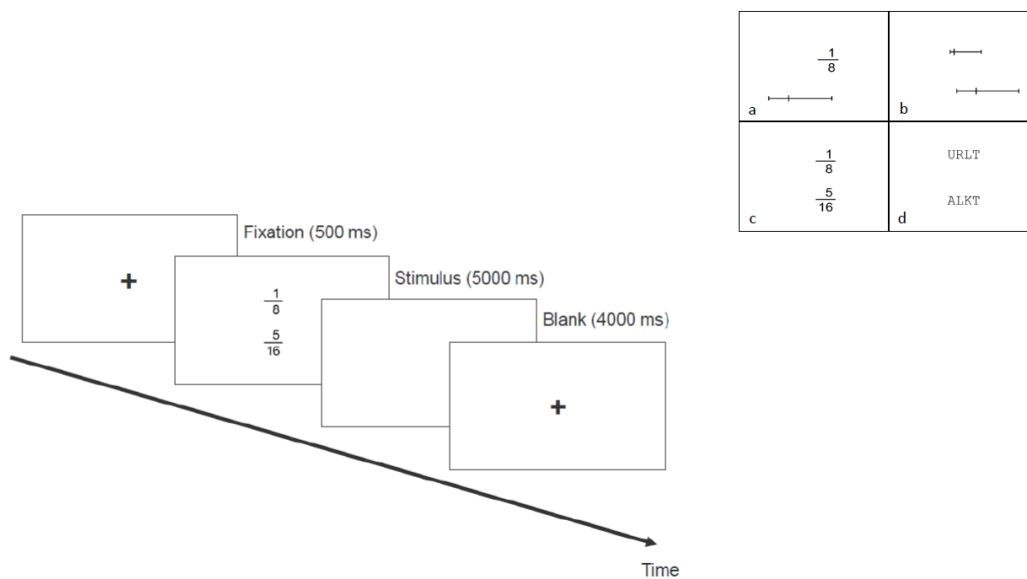
324 For the fMRI paradigm a block design was used, with four separate blocks for the three  
325 different tasks (i.e., fraction-line proportion matching task, line proportion comparison task and  
326 symbolic fraction magnitude comparison task). For each item in the fraction-line proportion  
327 matching task we presented a matching (i.e., magnitude of symbolic fraction and line proportion  
328 were identical) and a non-matching version (i.e., magnitude of symbolic fraction and line  
329 proportion were not identical) in the fraction-line proportion matching task. Therefore, this task  
330 took twice as long than both comparison tasks and was run in two blocks.

331 In the fraction-line proportion matching task, participants were shown a fraction and a  
332 line proportion (see Figure 3a). They had to indicate whether the fraction and the line proportion  
333 reflected the same magnitude or not (i.e., identical: left button, different: right button). In half  
334 of the items, magnitudes of fraction and line proportion matched. In the line proportion  
335 comparison task (see Figure 3b), participants were shown two line proportions and had to  
336 decide which proportion was the numerically larger one by pressing a corresponding response  
337 button (i.e., the right button when the upper and the left button when the lower line proportion  
338 was larger). Similar, in the symbolic fraction magnitude comparison task (see Figure 3c),  
339 participants were shown two fractions and had to decide which fraction was numerically larger  
340 again.

341 Each block consisted of 4 practice followed by 48 critical trials. In both sessions, half  
342 of the critical trials consisted of trained stimuli while the other half were untrained stimuli.  
343 Stimulus order was random for each participant and each session. Additionally, 22 trials of a

344 scrambled word task were randomly interspersed in each block of each condition to control for  
 345 eye-movements and to control that participants would stay focused. During these trials, two  
 346 strings of scrambled letters were shown on top of each other and participants had to decide  
 347 which of the strings would form a real word (see Figure 3d).

348 To prevent adaptation of the BOLD signal, 6000 ms pauses (white screen, RGB values  
 349 = 255 255 255) were randomly interspersed between trials. All stimuli were projected on a  
 350 screen outside the scanner and made visible to participants through a mirror mounted on the  
 351 head coil of the scanner. Foam pads were used to minimize head movements within the head  
 352 coil during fMRI acquisition. Stimuli were presented in black font against a white background  
 353 (RGB values = 255 255 255). The experiment was programmed using Presentation® v16.1  
 354 software (www.neurobs.com).



355  
 356 **Fig. 3: Experimental procedure with examples (upper right box) for the different tasks.** Example for a) the  
 357 fraction-line proportion matching task, b) the line proportion comparison task, c) the symbolic fraction magnitude  
 358 comparison task, and d) the scrambled letters control task. In this example the lower four letters can be unscrambled  
 359 to form the word “kalt” (German word for “cold”). The upper four letters cannot be formed to any word used in  
 360 German.  
 361

362 Each trial started with a fixation cross (500 ms), followed by the respective stimulus  
363 which appeared for 5000 ms or until participants responded. Subsequently, a blank screen  
364 appeared for 4000 ms followed by the beginning of a new trial (see Figure 3). Participants  
365 responded by pressing one of two MRI compatible response buttons with either their left or  
366 their right thumb. Participants were instructed to answer as fast and as accurately as possible.

367

## 368 **MRI and fMRI Acquisition**

369 A high-resolution T1-weighted anatomical scan was acquired by a 3T Siemens  
370 Magnetom Prisma MRI system (Siemens AG; Erlangen, Germany) equipped with a 64-channel  
371 head-neck matrix coil (TR = 2400 s, matrix =  $256 \times 256$ , 176 slices, voxel size =  $1.0 \times 1.0 \times$   
372  $1.0 \text{ mm}^3$ ; FOV = 256 mm, TE = 2.92 ms; flip angle =  $8^\circ$ ). The anatomical scan was always  
373 performed at the end of each experimental session.

374 Functional T2\*-weighted images were obtained using multiband gradient-echo Echo  
375 planar imaging (EPI; TR = 792 ms; TE = 30 ms; flip angle =  $58^\circ$ ; FOV = 192 mm,  $64 \times 64$   
376 matrix; 48 slices, voxel size =  $3.0 \times 3.0 \times 3.0 \text{ mm}^3$ ). Total scanning time was approximately 80  
377 minutes. A baseline (rest) condition was accomplished by including about 20% null events in  
378 the paradigm.

## 379 **Analysis**

### 380 **Preliminary Analysis**

381 Prior to the analysis of the behavioral and imaging data of the present study, possible  
382 differences between the two variants of NLE training (*gamified* vs. *non game-based* training)  
383 were examined both on the behavioral as well as the neuro-functional level. However, the  
384 behavioral analysis after the training showed no significant differences in reaction times or error  
385 rates for all three evaluation tasks performed in the scanner (i.e., symbolic fraction magnitude  
386 comparison, line proportion comparison, and fraction-line proportion matching task, all *p*-

387 values  $> .05$ , all  $F_s \leq 2.85$ ). In line with this, the analysis of imaging data revealed no significant  
388 difference between the two groups after the training for any of the three evaluation tasks at an  
389 uncorrected  $p$ -value of  $.001$  with a cluster size of  $k = 10$  voxels. Because neither behavioral nor  
390 neurofunctional differences were observed for the two trainings, we decided to merge both  
391 training groups in order to investigate fraction learning across groups with higher statistical  
392 power.

393

394 Behavioral analysis

395 *Training and flow data*

396 To evaluate learning outcomes over the five training time points and associations with  
397 participants' flow experience during learning we used a latent growth linear mixed-effects  
398 model over the five training time points. Regarding dependent variables, we were interested in  
399 the mean percentage absolute estimation error (PAE; [55]) and the number of attempts  
400 participants needed to estimate a given fraction correctly. Models were fitted in R using 'lmer'  
401 from the 'lme4' package [56]. To provide  $p$ -values we used the 'summary' function of the  
402 "lmerTest" R package [57]. Summary statistics were extracted via the 'analyze' function of  
403 'psycho' [58].

404 *Number line estimation task*

405 Mean PAE (cf. [55]) was calculated to reflect performance in the number line estimation  
406 task at the two time points. Items without a response were not considered for analyses. To  
407 evaluate performance changes in PAE between the pre- vs. post-test, a linear mixed-effects  
408 model was fitted using 'lmer' from the "lme4" R package [56]. Again, to provide  $p$ -values we  
409 used the 'summary' function of the "lmerTest" R package [57]. Additionally, summary  
410 statistics were also extracted via the 'analyze' function of "psycho" R package [58].

411



412 *Behavioral fMRI data*

413 For the analysis of the behavioral fMRI data we evaluated reaction times and accuracy  
414 as dependent variables. Three separate linear mixed-effects models were run to analyze reaction  
415 times for the three different evaluation tasks. Items without a response and items answered  
416 incorrectly were not considered for reaction time analyses. To analyze error rates and to include  
417 participants' individual performance as random effect, we ran three separate generalized linear  
418 mixed-effects models (GLMM) fitted by using the R package 'lme4' [56]. In the GLMMs we  
419 assumed a binomial error distribution and used the logit as the link function. For both types of  
420 analysis, we provided *p*-values with the 'summary' function of the "lmerTest" R package [57].

421

422 *Imaging analysis*

423 fMRI data analyses were performed using SPM12 (<http://www.fil.ion.ucl.ac.uk/spm>).  
424 Images were motion corrected and realigned to each participant's mean image. The mean image  
425 was co-registered with the anatomical whole-brain volume. Imaging data was then normalized  
426 into standard stereotaxic MNI space (Montreal Neurological Institute, McGill University,  
427 Montreal, Canada). Images were resampled every 2.5 mm using 4th degree spline interpolation  
428 to obtain isovoxels and then smoothed with a 5 mm full-width at half-maximum (FWHM)  
429 Gaussian kernel to accommodate inter-subject variation in brain anatomy and to increase signal-  
430 to-noise ratio in the images. Data were high-pass filtered (128 s) to remove low-frequency noise  
431 components and corrected for autocorrelation assuming an AR(1) process. Brain activity was  
432 convolved over all experimental trials with the canonical hemodynamic response function  
433 (HRF) and its first time derivative.

434 For the first level analysis, pre- and post-fMRI training sessions were combined on the  
435 subject level in a generalized linear model (GLM). For each participant, we considered the two  
436 factors item-set (trained vs. untrained) and session (pre vs post). This resulted in four  
437 experimental conditions: trained pre (T1), trained post (T2), untrained pre (UT1) and untrained

438 post (UT2). Additionally, to evaluate the influence of fraction problem difficulty we included  
439 the covariate numerical distance between to-be-compared fractions as a parametric regressor in  
440 the first level analysis. As a control variable the scrambled word problems (hereafter referred to  
441 as “words”) were also included in the first level. Finally, the six movement parameters from  
442 preprocessing were entered into the model to capture signal variations due to head motion.

443 These images then entered the second level random-effects group analysis using a  
444 flexible factorial design. The SPM Anatomy Toolbox [59], available for all published  
445 cytoarchitectonic maps ([http://www.fz-juelich.de/ime/spm\\_anatomy\\_toolbox](http://www.fz-juelich.de/ime/spm_anatomy_toolbox)), was used for  
446 anatomical localization of effects where applicable. For areas not yet implemented, the  
447 anatomical automatic labelling tool (AAL) in SPM12 ([http://www.cyceron.fr/web/aal](http://www.cyceron.fr/web/aal_anatomical_automatic_labeling.html)  
448 [anatomical\\_automatic\\_labeling.html](http://www.cyceron.fr/web/aal_anatomical_automatic_labeling.html)) was used.

449 All contrasts calculated reflect the parametric modulation of the fMRI signal by the  
450 numerical distance between two proportions presented (distance effect). Simple contrasts  
451 (distance effect in each notation) were family-wise error corrected at  $p < .05$  at the whole-brain  
452 level with a cluster size of  $k = 10$  voxels. Complex contrasts comparing two distance effects  
453 (e.g., distance effect in fractions after training vs. distance effect prior to the training) were  
454 thresholded at an uncorrected  $p$ -value of  $< .001$  at the voxel level with a cluster size of  $k = 10$   
455 voxels and were reported when they remained significant following family-wise error  
456 correction (FWE) at the cluster-level with  $p_{\text{cluster-corr}} < .05$ .

457

## 458 **Results**

### 459 **Behavioral Results**

#### 460 **Training and Flow Experience**

##### 461 *Percentage absolute estimation error*

462 Differences in PAE and possible associations with flow experience over time were  
463 analyzed using a latent growth linear mixed effect model, predicting PAE by flow experience  
464 and time (i.e., five training time points) as fixed factors while also including a random intercept

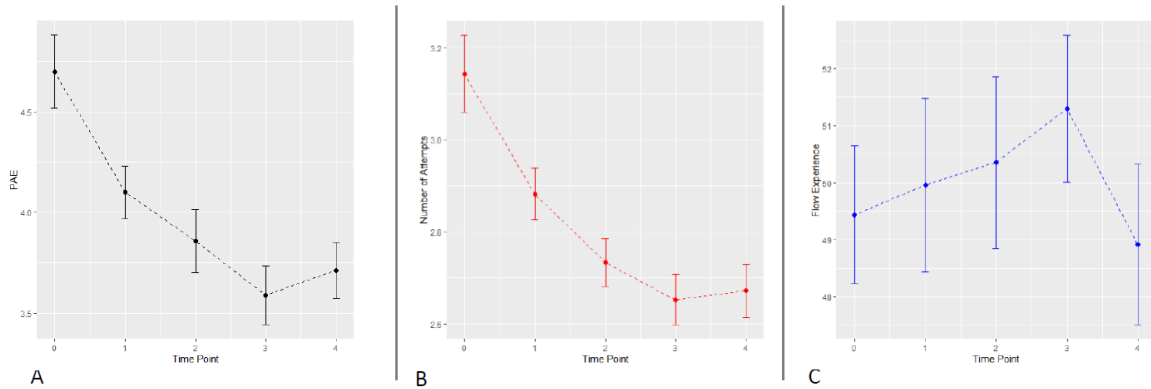
465 to account for participants' individual differences in prior knowledge. The model explained a  
466 significant proportion of variance in PAE ( $R^2 = 75.22\%$ ; fixed effects:  $R^2 = 12.97\%$ ) and showed  
467 that PAE [ $\beta = -0.34$ ,  $SE = 0.15$ ,  $t(193) = -2.29$ ,  $p < .05$ ] significantly improved over time (see  
468 Figure 4A for PAE changes over time). Moreover, the fixed effect of flow [ $\beta = -0.02$ ,  $SE =$   
469  $0.01$ ,  $t(198) = -3.13$ ,  $p < .01$ ] was significant, indicating that flow experience changed over  
470 time (see Figure 4C for general changes in flow experience over time). The interaction between  
471 time and flow experience was not significant [ $t(193) = 0.62$ ,  $p = .54$ ].

472

### 473 *Number of attempts*

474 Identical to the first analysis, differences in the number of attempts participants needed  
475 to estimate the given fraction during the NLE training correctly over time and possible  
476 associations with flow experience were analyzed using again a latent growth linear mixed effect  
477 model. Number of attempts needed was predicted by participants' flow experience and time  
478 (i.e., five training time points) as fixed factors. Again, we included a random intercept to  
479 account for participants individual differences in prior knowledge. The model explained a  
480 significant proportion of variance in the number of attempts needed to estimate the given  
481 fraction correctly ( $R^2 = 71.24\%$ ; fixed effects:  $R^2 = 16.16\%$ ) and showed that the number of  
482 attempts significantly decreased over time [ $\beta = -0.18$ ,  $SE = 0.07$ ,  $t(194) = -2.70$ ,  $p < .01$ ; see  
483 Figure 4B for number of attempts changes over time]. Within this model the fixed effect of flow  
484 was significant [ $\beta = -0.01$ ,  $SE = 0.06$ ,  $t(201) = -3.40$ ,  $p < .001$ ], indicating that flow experience  
485 changed over time (see Figure 4C for general changes in flow experience over time). Again,  
486 the interaction between time and flow experience was not significant [ $t(194) = 0.95$ ,  $p = .34$ ].

487



488  
 489 **Fig. 4:** Improvement of PAE (A), number of attempts needed (B), and changes in flow experience (C) over the  
 490 training period (i.e., five time points). Error bars indicate standard errors (SEM).  
 491  
 492

493 **Number line estimation task**

494 To investigate whether PAE for the paper pencil-based NLE tasks differed between pre-  
 495 and post-test we ran another linear mixed effect model. We defined the two time points (pre vs.  
 496 post) as fixed effect and included a random intercept for subjects to account for individual  
 497 variability. The overall model predicting differences in PAE on the two number line estimation  
 498 tests (pre vs. post) explained a significant proportion of variance ( $R^2 = 63.34\%$ , in which the  
 499 fixed effects explained  $R^2 = 5.56\%$  of the variance). The effect of session was significant [ $\beta =$   
 500  $-0.01$ ,  $SE = 0.00$ ,  $t(47) = -3.80$ ,  $p < .001$ ] indicating that performance was better on the posttest  
 501 than the pretest (reflected by smaller estimation errors).

502  
 503 **Behavioral fMRI results**

504 *Fraction-line proportion matching task*

505 Performance changes in reaction time between pre- and post fMRI sessions for the  
 506 *fraction-line proportion matching task* were evaluated by a linear mixed effect model. We  
 507 defined session (pre vs. post), itemset used (trained vs. untrained), and numerical distance (i.e.,  
 508 only for the non-matching items) as fixed effects. Moreover, interactions between session and  
 509 item set as well as session and numerical distance were also included as fixed effects to evaluate

510 whether the possible influence of item set or numerical distance on participants reaction time  
511 changed from pre- to post-test. Finally, we included a random intercept to account for  
512 participants individual differences in prior knowledge.

513 The model explained significant proportions of variance on participants reaction times  
514 ( $R^2 = 26.18\%$ ; fixed effects:  $R^2 = 9.74\%$ ) and showed that reaction times significantly improved  
515 from pre- to post-test [ $\beta = -260.85$ ,  $SE = 89.37$ ,  $t(1068) = -2.92$ ,  $p < .01$ ]. Additionally, the  
516 fixed effect of numerical distance was significant [ $\beta = -1306.84$ ,  $SE = 210.59$ ,  $t(1069) = -6.21$ ,  
517  $p < .001$ ], indicating that reaction times improved more strongly for increasing distances.  
518 Neither the fixed effect of itemset nor the interactions were significant [all  $t \leq 0.48$ , all  $p_s > .05$ ].

519 For the evaluation of performance changes in accuracy between pre- and post-test, we  
520 ran a generalized linear mixed-effects model, by using logit as the link function and assuming  
521 a binomial distribution of the error rates. To avoid overfitting of the model we only included  
522 session, itemset used and numerical distance as fixed effects and a random intercept accounting  
523 for individual differences in prior knowledge. The model revealed a significant fixed effect of  
524 numerical distance [ $z = -6.86$ ,  $p < .001$ ], indicating that increasing distances between fractions  
525 and line proportions led to less errors. The fixed effects of session and trained item were not  
526 significant [all  $z \leq -0.18$ , all  $p_s > .05$ ].

### 527 *Line proportion comparison task*

528 Identical to the analysis of the *fraction-line proportion matching task* we ran the same  
529 linear mixed-effects model (including session, itemset used, numerical distance, as well as the  
530 interaction session and item set and the interaction session and numerical distance as fixed  
531 effects and a random intercept to account for individual differences in prior knowledge) to  
532 investigate possible performance changes in reaction time on the *line proportion comparison*  
533 *task* between pre- and post fMRI session. The model explained a significant proportion of  
534 variance on participants reaction times ( $R^2 = 42.71\%$ ; fixed effects:  $R^2 = 8.57\%$ ) and showed

535 that reaction times significantly improved from pre to post-test [ $\beta = -279.33$ ,  $SE = 118.19$ ,  
536  $t(507) = -2.36$ ,  $p < .05$ ]. Additionally, the fixed effect of numerical distance was significant [ $\beta$   
537  $= -1686.98$ ,  $SE = 295.05$ ,  $t(510) = -5.72$ ,  $p < .001$ ], indicating that reaction times got  
538 significantly faster with increasing distances. Neither the fixed effect of itemset nor the  
539 interactions were significant [all  $t \leq 0.35$ , all  $p_s > .05$ ].

540         Again, to investigate possible performance changes in accuracy for the *line proportion*  
541 *comparison task* between pre- and post-test, we ran a generalized linear mixed-effects model,  
542 using logit as the link function and assumed a binomial distribution of the error rates. To avoid  
543 overfitting of the model we only included session, itemset used and numerical distance as fixed  
544 effects and a random intercept accounting for individual differences in prior knowledge.  
545 Analyzing error rates for the line proportion comparison task revealed a significant fixed effect  
546 of numerical distance [ $z = -7.5$ ,  $p < .001$ ], indicating that increasing distances between line  
547 proportions led to less errors. However, there were no significant differences for session and  
548 item set [all  $z \leq -1.10$ , all  $p_s > .05$ ].

#### 549 *Symbolic fraction magnitude comparison task*

550         Finally, possible performance changes in reaction time for the *symbolic fraction*  
551 *magnitude comparison task* between pre- and post-test were tested identical to the prior two  
552 evaluation tasks. We again ran a linear mixed-effects model with session, itemset used,  
553 numerical distance as well as the interaction session and item set and the interaction session and  
554 numerical distance as fixed effects and a random intercept accounting for individual differences  
555 in prior knowledge. Interestingly, in this model only the fixed effect for numerical distance was  
556 significant [ $\beta = -1282.76$ ,  $SE = 337.30$ ,  $t(499) = -3.80$ ,  $p < .001$ ], indicating that reaction times  
557 were significantly faster for increasing distances. Neither the fixed effects of session and itemset  
558 nor the interactions were significant [all  $t \leq 0.43$ , all  $p_s > .05$ ].

559 Possible performance changes in accuracy for the *symbolic fraction magnitude*  
560 *comparison* task between pre- and post-test, were evaluated again by a generalized linear  
561 mixed-effects model, using logit as the link function and assuming a binomial distribution of  
562 the error rates. To avoid overfitting of the model we only included session, itemset used and  
563 numerical distance as fixed effects and a random intercept accounting for individual differences  
564 in prior knowledge. Thus, analyzing error rates for the fraction task revealed a significant fixed  
565 effect of session [ $z = -2.38, p < .05$ ], indicating that participants performed better in the post-  
566 test compared to the pre-test. Additionally, the fixed effect of numerical distance was significant  
567 [ $z = -9.17, p < .001$ ], indicating that increasing distances between fractions led to less errors  
568 when comparing two fractions. However, there was no significant difference for item set [ $z =$   
569  $0.282, p = .079$ ].

570

## 571 **Imaging results**

572 Distance effect before the training

573 *Fraction-line proportion matching task*

574 Processing of smaller numerical distances between fractions and line proportions in the  
575 fraction-line proportion matching task was associated with stronger magnitude-specific fMRI  
576 signal before the training in a bilateral fronto-parietal neural network including areas in the  
577 intraparietal cortex (hIP3), the superior parietal cortex (SPL), the inferior frontal gyrus (Areas  
578 44 and 45), bilateral inferior temporal gyri as well as bilateral insula. Further activated clusters  
579 were found in the bilateral middle frontal gyri as well as right-hemispheric subcortical areas  
580 such as thalamus and caudate nucleus as well as the cerebellum (see Table 1, Figure 5A and C  
581 depicted in red color).

582 *Line proportion comparison task*

583 Processing of smaller numerical distance between to-be-compared line proportions  
584 modulated the fMRI signal before the training in a right-hemispheric fronto-parietal network  
585 centered around the right intraparietal sulcus (hIP3). Smaller numerical distance between  
586 proportions led to stronger activation in the right IPS and the right anterior IPS reflecting  
587 fraction magnitude processing. Additionally, there was a significant cluster of activation in the  
588 right inferior frontal gyrus (see Table 1, Figure 5B and C depicted in golden color).

589 *Symbolic fraction magnitude comparison task*

590 Activation in no cluster of voxels was modulated significantly by numerical distance at  
591 the given threshold in the symbolic fraction magnitude comparison task before the training.

592

593 Importantly, there were no significant differences in brain activation observed between  
594 the two stimulus sets: when comparing the two sets before the training (set to be trained vs.  
595 untrained set) in any of the three conditions (i.e., fraction-line proportion matching task, line  
596 proportion comparison task, symbolic fraction magnitude comparison task).

597

598

599 **Table 1:** Distance effect for line proportion comparison and fraction-line proportion matching task at a  
600 familywise error corrected  $p < .05$ , cluster size  $k = 10$  at the whole brain level.

601

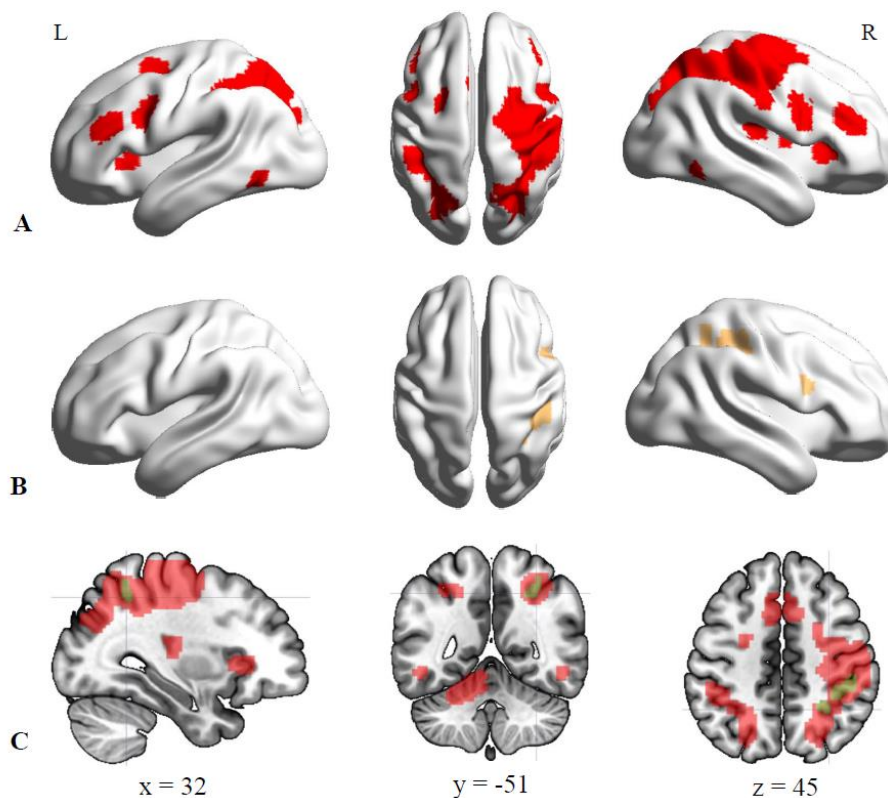
Contrast	Brain region	MNI (x, y, z)	Cluster size	$t$
Distance effect	RH anterior intraparietal sulcus (hIP2)	43 -36 47	61	5.63
<i>Line proportion</i>	RH intraparietal sulcus (hIP3)	33 -52 53	12	5.04
	RH inferior frontal gyrus (44)	51 8 23	15	5.61
Distance effect	RH precentral gyrus	38 -20 55	2195	11.49
<i>Matching task</i>	RH intraparietal sulcus (hIP3)	26 -57 55		7.15
	RH superior parietal lobe (SPL)	18 -60 58		5.46
	LH intraparietal sulcus (hIP3)	-30 -58 42	531	7.29



LH superior parietal lobe (SPL)	-25	-57	53		7.01
LH inferior frontal gyrus (IFG 44)	-42	5	30	167	7.71
RH inferior frontal gyrus (IFG 44)	53	8	28	179	7.38
LH inferior frontal gyrus (IFG 45)	-40	28	23	85	6.77
LH middle frontal gyrus	-20	6	55	58	6.25
RH middle frontal gyrus	41	41	18	85	5.91
LH posterior medial frontal gyrus	-7	8	58	271	7.24
RH inferior temporal gyrus	51	-52	-10	22	5.91
LH inferior temporal gyrus	-50	-57	-10	30	5.79
RH thalamus	18	-22	10	68	7.57
RH caudate nucleus	8	16	3	20	6.24
RH insula	33	-20	15	86	6.76
RH insula	31	28	3	54	6.08
RH insula	41	1	10	13	5.75
LH insula	-35	18	3	17	5.55
LH cerebellum	-17	-52	-23	306	10.56

602 Abbreviations: LH – left hemisphere; RH – right hemisphere; MNI – montreal neurological institute).

603



604

605 **Figure 5:** Effect of overall magnitude processing in the fractions-lines matching task (Panels A and C: red) and  
606 the lines proportion comparison task (Panels B and C: gold) as reflected by the distance effect.

607

608 Training and transfer effects

609 *Line proportion comparison / fraction-line proportion matching task*

610 When comparing the distance effect after the training to the distance effect before the  
611 training, no suprathreshold clusters of activation were observed for the line proportion  
612 comparison or in the fraction-line proportion matching task.

613 *Symbolic fraction magnitude comparison task*

614 Direct comparison of the distance effect after the training to the distance effect before  
615 the training in the symbolic fraction magnitude comparison task revealed significant increased  
616 activation differences in a bilateral fronto-parietal network centered around the intraparietal  
617 sulcus (hIP3; see Table 2, Figure 6). Further clusters of significant increased activation  
618 differences were observed in the right superior parietal lobe (SPL) and the left inferior parietal  
619 lobe (PFt), the right fusiform gyrus, the bilateral frontal cortex and the left thalamus.

620 When comparing distance effects observed for the trained item set to those observed for  
621 untrained item set after the training, no suprathreshold clusters of activation were observed,  
622 indicating that the training effect was comparably strong for trained as well as untrained fraction  
623 items. This means that for fraction magnitude processing it seemed that the training effect  
624 generalized to untrained items after the NLE training.

625

626 **Table 2:** Effect of training for processing of overall fraction magnitude as reflected by the distance  
627 effect for the symbolic fraction magnitude comparison task.

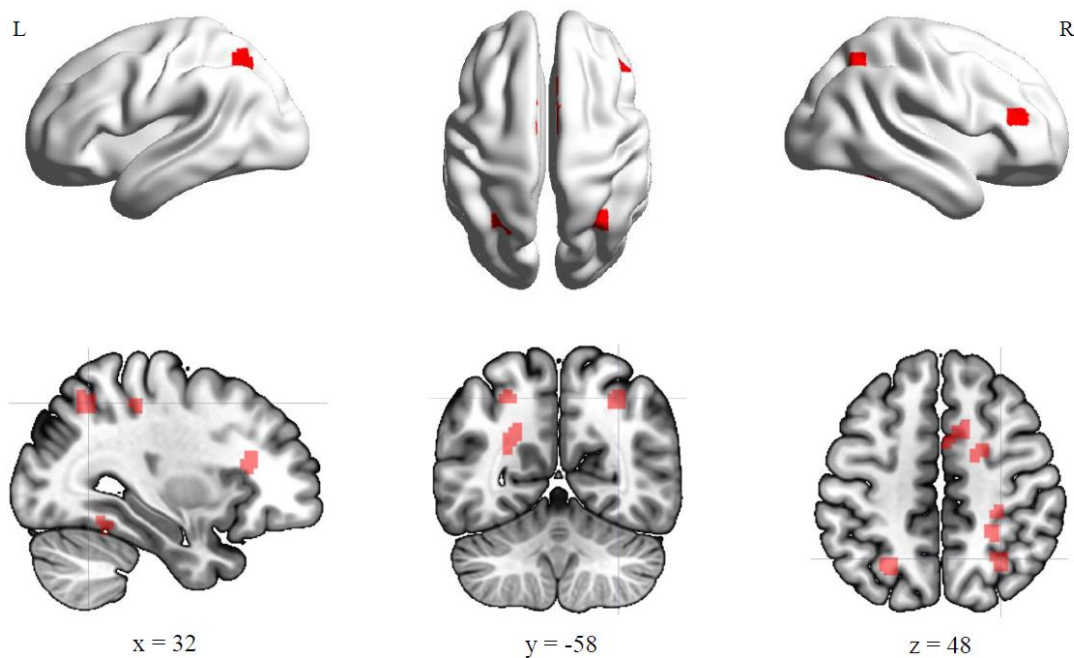
---

Contrast	Brain region	MNI (x, y, z)	Cluster size	<i>t</i>
----------	--------------	---------------	--------------	----------

---

Fractions post vs. pre training	RH superior parietal lobe (SPL)	26	-45	45	59	4.42
	RH intraparietal sulcus (hIP3)	31	-60	48	21	3.70
	LH intraparietal sulcus (hIP3)	-27	-62	50	15	3.53
	LH supramarginal gyrus (PFt)	-34	-28	33	55	4.56
	LH middle frontal gyrus	26	1	40	80	5.36
	RH posterior medial frontal gyrus	13	11	53	133	4.08
	RH inferior frontal gyrus (IFG)	41	36	15	45	4.05
	RH fusiform gyrus	36	-47	-18	15	3.51
	LH thalamus	-17	-5	3	18	5.61
	LH cuneus	-25	-57	28	15	4.07

628 Abbreviations: LH – left hemisphere; RH – right hemisphere; MNI – montreal neurological institute).



629

630 **Figure 6:** Effect of training for processing of overall fraction magnitude as reflected by the distance effect for the  
631 symbolic fraction magnitude comparison task ( $p_{\text{cluster-corr}} < .05$ ,  $k = 10$  voxels).

632

633

634

## Discussion

635           The aim of the present study was to investigate potential neuro-functional changes of  
636 brain activation in adult participants through a five-day consecutive NLE training on fraction  
637 magnitude. While there already exist a number of studies indicating the effectiveness of NLE  
638 training on the behavioral level [20–22,60] and some few studies investigating the neural  
639 correlates of fraction and proportion processing [41,42,44,45], little is still known about the  
640 neural correlates of fraction learning.

641           Evaluation tasks included symbolic fraction and non-symbolic line proportion  
642 magnitude comparison tasks and a fraction-line proportion matching task. Additionally,  
643 learning trajectories over the five time points of the NLE training, participants' corresponding  
644 flow experience as well as a paper pencil pre- post NLE task were evaluated. In the following,  
645 behavioral as well as neurofunctional results for these measures will be discussed in more detail.

646

#### 647 *The interplay between training performance and flow experience*

648           Flow experience has been described as an optimal state for intrinsically motivated  
649 learning, which helps focus on the given task and can lead to improved performance [31].  
650 Another explanation for optimal learning was first described by Yerkes and Dodson in 1908  
651 and refers to the relationship between arousal and performance (i.e., Yerkes-Dodson-Law;  
652 [61]). This law states that learners' optimal performance is achieved on a medium level of  
653 arousal reflected by an inverted U-shape relation between the respective parameters.  
654 Transferred to flow experience this inverted U-shape suggests that a balance between cognitive  
655 demands of the task at hand and individual skill level is the basis for best possible flow  
656 experience. Thus, when skills of a learner are too poor for the demands of a given task – for  
657 instance at the beginning of a training – flow experience may be rather low. The same is true  
658 when skills of a learner are too advanced for a given task. Both non-optimal states can lead to  
659 boredom and/or frustration, and reduced flow experience. In turn, this might interfere with  
660 learning of the given task [62].

661 In line with these assumptions, our results for training performance and flow experience  
662 over the five day NLE training indicated that at the beginning of the training participants' flow  
663 experience was significantly lower and PAE as well as number of attempts needed to solve a  
664 trial successfully were significantly higher as compared to later training days. This possibly  
665 reflects an imbalance between task demands and individual skills. PAE and number of attempts  
666 needed on the first training day suggest that participants' ability on the task seemed to be low  
667 in the beginning. With each training day passing flow experience increased while at the same  
668 time PAE and number of attempts needed decreased. Thus, participants experienced a more and  
669 more optimal learning situation in which demands of the task and individual skill level were in  
670 balance.

671 Interestingly, on the last training day participants' flow experience dramatically  
672 decreased again to values lower than on the first training day. This was accompanied by slight  
673 increases of PAE and number of attempts needed. Thus, we assume that participants peak of  
674 performance was already achieved at the 4<sup>th</sup> training day. We can only speculate whether this  
675 was caused either by boredom or the fact that fraction magnitudes could not be estimated more  
676 accurately by our participants after 4 days of training. Importantly, however, this decrease in  
677 flow experience was not associated with a general decrease in training performance. Moreover,  
678 we observed that participants significantly improved in the paper pencil NLE task from pre to  
679 post session.

#### 680 *Transfer effects in fraction and proportion learning*

681 Behavioral data indicated significant performance improvements for all three tasks.  
682 Importantly, these improvements did not differ between trained and untrained items, indicating  
683 transfer effects of the training to untrained items. Additionally, neurofunctional data showed  
684 similar results: before the training no significant differences in brain activation were observed  
685 between the two stimulus sets (trained and untrained set) for all three evaluation tasks. After  
686 the training, again, no suprathreshold clusters of activation were observed when comparing

687 trained and untrained items for all three evaluation tasks. Thus, indicating that the training effect  
688 was comparably strong for trained as well as untrained items and seemed to generalize to  
689 untrained items after the NLE training.

690 Moreover, as discussed in more detail below, for the case of the fraction magnitude  
691 comparison task we think brain activation associated with numerical distance after the NLE  
692 training indicates that overall symbolic fraction magnitude was not automatically processed  
693 before training.

694 However, one might argue that our results were elicited by the applied drill-like training  
695 approach. In the literature, this is often used to investigate numerical learning in terms of  
696 arithmetic fact learning [63,64]. Nevertheless, we are confident that participants did not just  
697 learn specific fraction magnitudes by heart for at least two reasons: First, if fraction magnitudes  
698 were learned by drill no transfer effect from trained to untrained items should be evident neither  
699 on the behavioral nor on the neural level. Moreover, fractions that are learned by heart should  
700 not show a numerical distance effect especially for untrained items because their overall  
701 magnitude should not be processed. In our study induced automated magnitude activation was  
702 also present for untrained items. Thus, it is unlikely that our training supported fact learning but  
703 rather improved magnitude representation.

704 Second, previous literature reported different neural correlates for arithmetic fact  
705 learning than observed in the current study. In particular, learning arithmetic facts is associated  
706 with a shift from bilateral fronto-parietal processing around the IPS to a primarily left  
707 hemispheric network in the medial temporary lobe (MTL) involving the hippocampus (cf. [64–  
708 66]). However, in the present study, we rather observed a shift towards more activation within  
709 the fronto-parietal network of magnitude processing [67–69] – thus, indicating more explicit  
710 processing of overall fraction magnitude and not fact retrieval after the training.

711 This supports the notion that the training indeed resulted in a general conceptual  
712 improvement and automatization of fraction magnitude processing, in contrast to training fact  
713 retrieval of specific fraction magnitudes (cf. [64] for limited evidence of transfer effects in  
714 multiplication fact training).

715 *Differential neural activity patterns before training and possible implications*

716 Surprisingly, and not consistent with the previous literature on neural correlates of  
717 fraction processing brain activation before the training for the three tasks of interest revealed  
718 different activation patterns in the IPS associated with the numerical distance effect. For the  
719 non-symbolic line proportion comparison and the fraction-line proportion matching task we  
720 found significant neural activation patterns in the typical fronto-parietal network observed  
721 previously for proportion and fraction processing (cf. [40]). In particular, the line proportion  
722 comparison task led to increased activation in the right intraparietal sulcus, whereas the  
723 fraction-line matching task led to increased activation in the bilateral intraparietal sulcus. This  
724 is in line with research on brain activation for symbolic and non-symbolic magnitude processing  
725 [70].

726 Interestingly, for the symbolic fraction magnitude comparison task activation in no  
727 cluster of voxels was modulated significantly by numerical distance before the training. This  
728 was surprising as previous studies on the neural correlates of fraction magnitude processing  
729 consistently reported IPS activity to be modulated by numerical distance for fraction magnitude  
730 processing. Importantly, the presented magnitudes did not differ between the three evaluation  
731 tasks. This means, that participants of our study were generally able to process the presented  
732 magnitudes. However, access to symbolic fraction magnitudes during the respective fraction  
733 comparison task might be reduced probably because of the bipartite nature of fractions [10].  
734 Thus, fractions are more difficult to compare than for instance line proportions.

735           Moreover, our fraction items used in this study were more complex because of two  
736 reasons: i) Our fraction pairs did not involve fractions with common components. In this case,  
737 reasoning about the natural number components alone (i.e., processing numerators and  
738 denominators separately) might often not help to find the right solution when comparing two  
739 fractions without common components. However, we used fractions without common  
740 components as we wanted to specifically investigate and promote overall fraction magnitude  
741 processing. For instance, [71] found that mathematic experts showed a distance effect for  
742 overall fraction magnitude while comparing fractions without common components but not for  
743 fraction pairs with common components. Moreover, comparing fraction pairs with common  
744 components is typically susceptible to what has been called the natural number bias [72]. ii)  
745 We used fractions with numerators and denominators ranging from 1 to 30. This had two major  
746 reasons: a) the difficulty to match two item sets of fraction pairs on both overall numerical  
747 distance and problem size including only fractions without common components and b) the fact  
748 that we wanted to make sure that fractions were rather unfamiliar to participants to be able to  
749 investigate fraction learning on a neural level. Thus, participants may not have had a specific  
750 representation of the magnitude of the presented fractions prior to our study.

#### 751 *Training induced distance effect in the intraparietal sulcus for fraction magnitude processing*

752           In line with previous results of training studies, imaging results after the training showed  
753 that the processing of overall symbolic fraction magnitude was improved. In particular, distance  
754 related neural activation for symbolic fraction processing became significantly stronger from  
755 pre- to post-test in the bilateral intraparietal sulcus. This may indicate that our NLE training  
756 helped to establish easier access to the representation of overall fraction magnitude. These  
757 results are inconsistent with a NLE training study on natural numbers with pre-post-test fMRI  
758 comparing children with and without developmental dyscalculia [73]. After the NLE training  
759 both groups showed decreased activation of brain areas involved in number magnitude



760 processing (for instance bilateral middle and superior frontal gyrus and left intraparietal sulcus).  
761 The authors interpreted these results as reflection of more automatized processing of numerical  
762 magnitude after their training.

763 In this context, however, there is inconsistent evidence in the literature about the  
764 distance effect as an indicator for better numerical/ mathematical performance (see Moeller et  
765 al., 2011). On the one hand, there are studies showing that a larger distance effect was associated  
766 with poorer numerical/ mathematical performance [74,75]. On the other hand, there also are  
767 studies observing that a more pronounced distance effect was associated with better numerical/  
768 mathematical performance [50,76]. To accommodate these inconsistent lines of evidence, [50]  
769 suggested that the relation between the size of the distance effect and mathematical performance  
770 might not be linear but curvilinear instead. In particular, these authors suggested that the size  
771 of the distance effect is depended on two factors: i) automated access to processed magnitudes  
772 decreases the distance effect whereas ii) increasing task complexity may increase the distance  
773 effect while processing magnitudes.

774 In line with this argument, it needs to be noted that we do not necessarily think that a  
775 larger distance effect indicates better number/ fraction magnitude processing. However, in the  
776 present study the increase of the distance effect on a neural level in symbolic fraction magnitude  
777 comparison might nevertheless indicate more automatic access to overall fraction magnitude as  
778 task complexity was very high. Moreover, due to the bipartite nature of fractions [10] overall  
779 fraction magnitude may not have been as automatically activated prior to the NLE training.  
780 Thus, after the NLE training our participants may have built a more coherent fraction magnitude  
781 representation reflected by a larger distance effect. In line with the hypothetical curvilinear  
782 model by [50] we think that the distance effect might decrease again after having established  
783 the magnitude representation with further training.

784           Moreover, no significant differences between pre- and post-fMRI were found for the  
785 line proportion comparison and fraction-line proportion matching task. This may indicate that  
786 improvement towards more automated activation on a neural level was not achieved through  
787 training as activation related to the numerical distance effect was already there before training.  
788 Again, this might reflect that the bipartite nature of fractions might have hindered automated  
789 magnitude processing of the fractions before training. In turn, non-symbolic proportions are  
790 further reflected by visual-spatial aspects which may have helped process the actual relative  
791 magnitude expressed as compared to symbolic fractions for which this relation needs to be built  
792 by participants themselves. Thus, the NLE training might not have changed processing of line  
793 proportions and fraction-line matching significantly on the neural level as the respective relative  
794 magnitudes may have already been processed before the training due to facilitation by visual-  
795 spatial aspects of the presentation format.

796           Taken together, these results indicate that even well-educated adults benefitted from a  
797 NLE based training of fractions aimed at improving conceptual knowledge of fraction  
798 magnitude. Importantly, the training did not only induce significant training effects on the  
799 behavioral level but in particular also led to changes in brain activation associated with the  
800 processing of symbolic fraction magnitude. This indicates processes of neurofunctional  
801 plasticity in fraction learning. In the following, we will discuss implications of these results for  
802 education.

### 803 *Implications for education*

804           The final report of the National Mathematics Advisory Panel states that ‘one key  
805 mechanism linking conceptual and procedural knowledge is the ability to represent fractions on  
806 a number line’ (p. 28; [77]). However, fraction learning and understanding still is an educational  
807 challenge not only in the US but globally. The integrated theory of numerical development,  
808 postulates that one core basis of all (rational) numbers is their magnitudes and that these

809 magnitudes can be represented on a mental number line [13]. Students difficulties with fractions  
810 often arise from missing conceptual understanding, which among other things but not  
811 exclusively involves an understanding of their magnitudes [78]. Therefore, it is a crucial step  
812 for students to learn that fractions are numbers with magnitudes that can be represented on a  
813 number line as well.

814 Accordingly, the recommendation to use number lines as an instructional tool to foster  
815 conceptual understanding of fractions is given in different guidelines for educational practice  
816 in fraction teaching (e.g., *Teaching fractions* [78] or *Developing Effective Fraction Instruction*  
817 *for Kindergarten through 8<sup>th</sup> grade* [79]). This recommendation is supported by recent evidence  
818 from different intervention studies that used number lines as intervention tools and found  
819 significant improvements of children's performance and understanding of fractions [20–23,60].

820 With respect to educational practice, our results support the existing body of literature  
821 that processing of proportion and fraction magnitudes can be improved by NLE training.  
822 Moreover, to the best of our knowledge this is the first study indicating that such a training  
823 improves symbolic fraction processing as reflected by the numerical distance effect on a neural  
824 level. In particular, we argue that relative magnitude information of complex fractions may  
825 initially not be processed automatically within the IPS as indicated by the missing numerical  
826 distance related activation in the IPS before but significant activation associated with numerical  
827 distance after the training.

### 828 *Limitations*

829 When interpreting the results of the current study there are some limiting aspects that  
830 need to be considered. First of all, we are well aware that the current study is only a first step  
831 towards a better understanding of the underlying neural processes of fraction learning. In  
832 particular, this study investigated fraction learning on fractions more complex than those  
833 fractions first encounter in school. This was the case for two major reasons: i) Our fraction pairs

834 did not consist of fractions with common components, which limits the number of available  
835 fractions when only considering those with numerators and denominators ranging from 1 to 9.  
836 Therefore, we used fractions with numerators and denominators ranging from 1 to 30 allowing  
837 for proper matching of stimuli sets. ii) We ran our study with adult participants for whom we  
838 assumed that they should be more or less proficient with fractions with numerators and  
839 denominators up to 9. Thus, to be able to detect training effects we used more complex fractions.

840 As such, to investigate fraction learning more fundamentally, less complex fractions  
841 (i.e., with single-digit numerators and denominators or even unit fractions) should be used.  
842 Moreover, our study investigated fraction processing in adult participants. To focus more on  
843 the fundamentals of fraction learning the developing brain should be investigated. A first  
844 attempt, to investigate developmental differences in fraction magnitude processing on a neural  
845 level is a study by [80]. In this study, the authors applied 2 mA bilateral tDCS and found that  
846 adults and children benefitted differently during fraction processing by tDCS applied to  
847 different areas of stimulation (IPS vs. dorsolateral prefrontal cortex (DLPFC)). However,  
848 imaging studies on the neural correlates of proportion and fraction magnitude processing in  
849 children are still missing.

## 850 *Conclusion*

851

852 It is well known that fractions are difficult to learn and understand not only for children  
853 and students but even adults and teachers [7,81,82]. Therefore, the ability to foster and improve  
854 fraction knowledge is of high educational importance. Apart from beneficial effects of NLE  
855 training on the behavioral level, our study provides first insights into the neural correlates of  
856 fraction learning. In particular, we did not observe numerical magnitude to significantly  
857 modulate brain activation before the training for the processing of symbolic fractions. This  
858 might indicate that overall fraction magnitude is not yet processed automatically before  
859 training. Thus, through the training participants might have built up more automated processing

860 of overall fraction magnitude. As such, our results indicate a specific improvement of overall  
861 fraction magnitude processing through NLE training reflected on the neural level. This case of  
862 neuronal plasticity in fraction learning indicates neurofunctional changes elicited by training of  
863 educationally relevant content. Therefore, our study supports the importance of NLE trainings  
864 for fraction learning on a neurophysiological level.

865

866

867

### References

- 868 [1] D.H. Bailey, M.K. Hoard, L. Nugent, D.C. Geary, Competence with fractions predicts  
869 gains in mathematics achievement, *J. Exp. Child Psychol.* 113 (2012) 447–455.
- 870 [2] R.S. Siegler, G.J. Duncan, P.E. Davis-Kean, K. Duckworth, A. Claessens, M. Engel,  
871 M.I. Susperreguy, M. Chen, Early predictors of high school mathematics achievement,  
872 *Psychol. Sci.* 23 (2012) 691–697.
- 873 [3] J. Torbeyns, M. Schneider, Z. Xin, R.S. Siegler, Bridging the gap: Fraction  
874 understanding is central to mathematics achievement in students from three different  
875 continents, *Learn. Instr.* 37 (2015) 5–13.
- 876 [4] M.J. Behr, I. Wachsmuth, T.R. Post, R. Lesh, Order and equivalence of rational  
877 numbers: A clinical teaching experiment, *J. Res. Math. Educ.* (1984) 323–341.
- 878 [5] K. Hart, Fractions, *Math. Sch.* 10 (1981) 13–15.
- 879 [6] R.S. Siegler, A.A. Pyke, Developmental and individual differences in understanding of  
880 fractions., *Dev. Psychol.* 49 (2013) 1994.
- 881 [7] X. Vamvakoussi, S. Vosniadou, Understanding the structure of the set of rational  
882 numbers: A conceptual change approach, *Learn. Instr.* 14 (2004) 453–467.

- 883 [8] H. Lortie-Forgues, J. Tian, R.S. Siegler, Why is learning fraction and decimal  
884 arithmetic so difficult?, *Dev. Rev.* 38 (2015) 201–221.
- 885 [9] T.P. Carpenter, M.K. Corbitt, H.S. Kepner, M.M. Lindquist, R. Reys, Results of the  
886 second NAEP mathematics assessment: Secondary school, *Math. Teach.* 73 (1980)  
887 329–338.
- 888 [10] M. DeWolf, M.A. Grounds, M. Bassok, K.J. Holyoak, Magnitude comparison with  
889 different types of rational numbers., *J. Exp. Psychol. Hum. Percept. Perform.* 40 (2014)  
890 71.
- 891 [11] R.S. Siegler, Magnitude knowledge: The common core of numerical development,  
892 *Dev. Sci.* 19 (2016) 341–361.
- 893 [12] R.S. Siegler, H. Lortie-Forgues, An integrative theory of numerical development, *Child*  
894 *Dev. Perspect.* 8 (2014) 144–150.
- 895 [13] R.S. Siegler, C.A. Thompson, M. Schneider, An integrated theory of whole number  
896 and fractions development, *Cogn. Psychol.* 62 (2011) 273–296.
- 897 [14] R. Gersten, R.F. Schumacher, N.C. Jordan, Life on the number line: Routes to  
898 understanding fraction magnitude for students with difficulties learning mathematics, *J.*  
899 *Learn. Disabil.* 50 (2017) 655–657.
- 900 [15] S.M. Göbel, S. Shaki, M.H. Fischer, The cultural number line: a review of cultural and  
901 linguistic influences on the development of number processing, *J. Cross. Cult. Psychol.*  
902 42 (2011) 543–565.
- 903 [16] I. Berteletti, D. Lucangeli, M. Piazza, S. Dehaene, M. Zorzi, Numerical estimation in  
904 preschoolers., *Dev. Psychol.* 46 (2010) 545.
- 905 [17] D.C. Geary, M.K. Hoard, L. Nugent, J. Byrd-Craven, Development of number line

- 906 representations in children with mathematical learning disability, *Dev. Neuropsychol.*  
907 33 (2008) 277–299.
- 908 [18] R.S. Siegler, J.E. Opfer, The development of numerical estimation: Evidence for  
909 multiple representations of numerical quantity, *Psychol. Sci.* 14 (2003) 237–250.
- 910 [19] H.C. Barth, A.M. Paladino, The development of numerical estimation: Evidence  
911 against a representational shift, *Dev. Sci.* 14 (2011) 125–135.
- 912 [20] N. Hamdan, E.A. Gunderson, The number line is a critical spatial-numerical  
913 representation: Evidence from a fraction intervention., *Dev. Psychol.* 53 (2017) 587.
- 914 [21] C.A. Barbieri, J. Rodrigues, N. Dyson, N.C. Jordan, Improving fraction understanding  
915 in sixth graders with mathematics difficulties: Effects of a number line approach  
916 combined with cognitive learning strategies., *J. Educ. Psychol.* 112 (2020) 628.
- 917 [22] L.K. Fazio, C.A. Kennedy, R.S. Siegler, Improving children’s knowledge of fraction  
918 magnitudes, *PLoS One.* 11 (2016) e0165243.
- 919 [23] K. Kiili, K. Moeller, M. Ninaus, Evaluating the effectiveness of a game-based rational  
920 number training-In-game metrics as learning indicators, *Comput. Educ.* 120 (2018) 13–  
921 28.
- 922 [24] K. Kiili, K. Ojansuu, A. Lindstedt, M. Ninaus, Exploring the educational potential of a  
923 game-based math competition, *Int. J. Game-Based Learn.* 8 (2018) 14–28.
- 924 [25] J.L. Plass, U. Kaplan, Emotional design in digital media for learning, in: *Emot.*  
925 *Technol. Des. Learn.*, Elsevier, 2016: pp. 131–161.
- 926 [26] K. Kiili, Digital game-based learning: Towards an experiential gaming model, *Internet*  
927 *High. Educ.* 8 (2005) 13–24.
- 928 [27] M. Csikszentmihalyi, Beyond boredom and anxiety: Experiencing flow in work and

- 929 play, San Fr. (1975).
- 930 [28] K. Kiili, A. Lindstedt, M. Ninaus, Exploring characteristics of students' emotions, flow  
931 and motivation in a math game competition., in: GamiFIN, 2018: pp. 20–29.
- 932 [29] A. Landhäußer, J. Keller, Flow and its affective, cognitive, and performance-related  
933 consequences, in: *Adv. Flow Res.*, Springer, 2012: pp. 65–85.
- 934 [30] D.C. Moos, E. Marroquin, Multimedia, hypermedia, and hypertext: Motivation  
935 considered and reconsidered, *Comput. Human Behav.* 26 (2010) 265–276.
- 936 [31] M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience*. Perennial Modern  
937 Classics, (1990).
- 938 [32] M. Ulrich, J. Keller, K. Hoenig, C. Waller, G. Grön, Neural correlates of  
939 experimentally induced flow experiences, *Neuroimage*. 86 (2014) 194–202.
- 940 [33] M. Ulrich, J. Keller, G. Grön, Neural signatures of experimentally induced flow  
941 experiences identified in a typical fMRI block design with BOLD imaging, *Soc. Cogn.*  
942 *Affect. Neurosci.* 11 (2016) 496–507.
- 943 [34] A. Perttula, K. Kiili, A. Lindstedt, P. Tuomi, Flow experience in game based learning–  
944 a systematic literature review, *Int. J. Serious Games*. 4 (2017) 57–72.
- 945 [35] P. van Schaik, S. Martin, M. Vallance, Measuring flow experience in an immersive  
946 virtual environment for collaborative learningjcal\_455 350.. 365, (2011).
- 947 [36] U. Konradt, R. Filip, S. Hoffmann, Flow experience and positive affect during  
948 hypermedia learning, *Br. J. Educ. Technol.* 34 (2003) 309–327.
- 949 [37] L.-A. Ho, T.-H. Kuo, How can one amplify the effect of e-learning? An examination of  
950 high-tech employees' computer attitude and flow experience, *Comput. Human Behav.*  
951 26 (2010) 23–31.



- 952 [38] L.A. Custodero, Observable indicators of flow experience: A developmental  
953 perspective on musical engagement in young children from infancy to school age,  
954 *Music Educ. Res.* 7 (2005) 185–209.
- 955 [39] R. MacDonald, C. Byrne, L. Carlton, Creativity and flow in musical composition: An  
956 empirical investigation, *Psychol. Music.* 34 (2006) 292–306.
- 957 [40] S.N. Jacob, A. Nieder, Tuning to non-symbolic proportions in the human frontoparietal  
958 cortex, *Eur. J. Neurosci.* 30 (2009) 1432–1442.
- 959 [41] J. Mock, S. Huber, J. Bloechle, J.F. Dietrich, J. Bahnmueller, J. Rennig, E. Klein, K.  
960 Moeller, Magnitude processing of symbolic and non-symbolic proportions: an fMRI  
961 study, *Behav. Brain Funct.* 14 (2018) 9.
- 962 [42] J. Mock, S. Huber, J. Bloechle, J. Bahnmueller, K. Moeller, E. Klein, Processing  
963 symbolic and non-symbolic proportions: domain-specific numerical and domain-  
964 general processes in intraparietal cortex, *Brain Res.* 1714 (2019) 133–146.
- 965 [43] J. Cui, L. Li, M. Li, R. Siegler, X. Zhou, Middle temporal cortex is involved in  
966 processing fractions, *Neurosci. Lett.* (2020) 134901.
- 967 [44] M. DeWolf, J.N. Chiang, M. Bassok, K.J. Holyoak, M.M. Monti, Neural  
968 representations of magnitude for natural and rational numbers, *Neuroimage.* 141 (2016)  
969 304–312.
- 970 [45] A. Ischebeck, M. Schocke, M. Delazer, The processing and representation of fractions  
971 within the brain: An fMRI investigation, *Neuroimage.* 47 (2009) 403–413.
- 972 [46] S.N. Jacob, A. Nieder, Notation-independent representation of fractions in the human  
973 parietal cortex, *J. Neurosci.* 29 (2009) 4652–4657.
- 974 [47] R.S. Moyer, T.K. Landauer, Time required for judgements of numerical inequality,

- 975 Nature. 215 (1967) 1519–1520.
- 976 [48] M. Klabunde, M. Saggarr, K.M. Hustyi, R.G. Kelley, A.L. Reiss, S.S. Hall, Examining  
977 the neural correlates of emergent equivalence relations in fragile X syndrome,  
978 Psychiatry Res. Neuroimaging. 233 (2015) 373–379.
- 979 [49] M. Piazza, P. Pinel, D. Le Bihan, S. Dehaene, A magnitude code common to  
980 numerosities and number symbols in human intraparietal cortex, Neuron. 53 (2007)  
981 293–305.
- 982 [50] K. Moeller, S. Pixner, J. Zuber, L. Kaufmann, H.-C. Nuerk, Early place-value  
983 understanding as a precursor for later arithmetic performance—A longitudinal study on  
984 numerical development, Res. Dev. Disabil. 32 (2011) 1837–1851.
- 985 [51] F. Rheinberg, S. Engeser, R. Vollmeyer, Measuring components of flow: the Flow-  
986 Short-Scale, Proc. 1st Int. Posit. Psychol. Summit. (2002).
- 987 [52] M. Ninaus, K. Kiili, J. McMullen, K. Moeller, Assessing fraction knowledge by a  
988 digital game, Comput. Human Behav. 70 (2017) 197–206.
- 989 [53] M. Ninaus, K. Moeller, J. McMullen, K. Kiili, Acceptance of Game-Based Learning  
990 and Intrinsic Motivation as Predictors for Learning Success and Flow Experience.,  
991 (2017).
- 992 [54] & M. Ninaus, Kiili, Wortha, Motivationsprofile bei Verwendung eines Lernspiels zur  
993 Messung des Bruchverständnisses in der Schule - Eine latente Profilanalyse., Psychol.  
994 Erziehung Und Unterricht. (n.d.).
- 995 [55] J.L. Booth, R.S. Siegler, Developmental and individual differences in pure numerical  
996 estimation., Dev. Psychol. 42 (2006) 189.
- 997 [56] D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using

- 998 lme4, ArXiv Prepr. ArXiv1406.5823. (2014).
- 999 [57] A. Kuznetsova, P.B. Brockhoff, R.H.B. Christensen, lmerTest Package: Tests in Linear  
1000 Mixed Effects Models. R package version 3.0–1, J Stat Softw. (2018).
- 1001 [58] D. Makowski, The psycho package: an efficient and publishing-oriented workflow for  
1002 psychological science, J. Open Source Softw. 3 (2018) 470.
- 1003 [59] S.B. Eickhoff, K.E. Stephan, H. Mohlberg, C. Grefkes, G.R. Fink, K. Amunts, K.  
1004 Zilles, A new SPM toolbox for combining probabilistic cytoarchitectonic maps and  
1005 functional imaging data, Neuroimage. 25 (2005) 1325–1335.  
1006 <https://doi.org/10.1016/j.neuroimage.2004.12.034>.
- 1007 [60] E.A. Gunderson, N. Hamdan, L. Hildebrand, V. Bartek, Number line unidimensionality  
1008 is a critical feature for promoting fraction magnitude concepts, J. Exp. Child Psychol.  
1009 187 (2019) 104657.
- 1010 [61] R.M. Yerkes, J.D. Dodson, The relation of strength of stimulus to rapidity of habit-  
1011 formation, Punishm. Issues Exp. (1908) 27–41.
- 1012 [62] S. Greipl, K. Moeller, M. Ninaus, Potential and limits of game-based learning, Int. J.  
1013 Technol. Enhanc. Learn. (2020).
- 1014 [63] A. Ischebeck, L. Zamarian, M. Schocke, M. Delazer, Flexible transfer of knowledge in  
1015 mental arithmetic—An fMRI study, Neuroimage. 44 (2009) 1103–1112.
- 1016 [64] J. Bloechle, S. Huber, J. Bahnmüller, J. Rennig, K. Willmes, S. Cavdaroglu, K.  
1017 Moeller, E. Klein, Fact learning in complex arithmetic—the role of the angular gyrus  
1018 revisited, Hum. Brain Mapp. 37 (2016) 3061–3079.
- 1019 [65] V. Menon, Memory and cognitive control circuits in mathematical cognition and  
1020 learning, in: Prog. Brain Res., Elsevier, 2016: pp. 159–186.

- 1021 [66] E. Klein, K. Willmes, S.M. Bieck, J. Bloechle, K. Moeller, White matter neuro-  
1022 plasticity in mental arithmetic: changes in hippocampal connectivity following  
1023 arithmetic drill training, *Cortex*. 114 (2019) 115–123.
- 1024 [67] M. Amalric, S. Dehaene, Origins of the brain networks for advanced mathematics in  
1025 expert mathematicians, *Proc. Natl. Acad. Sci.* 113 (2016) 4909–4917.
- 1026 [68] M. Amalric, S. Dehaene, Cortical circuits for mathematical knowledge: evidence for a  
1027 major subdivision within the brain’s semantic networks, *Philos. Trans. R. Soc. B Biol.*  
1028 *Sci.* 373 (2018) 20160515.
- 1029 [69] M. Amalric, S. Dehaene, A distinct cortical network for mathematical knowledge in the  
1030 human brain, *Neuroimage*. 189 (2019) 19–31.
- 1031 [70] H.M. Sokolowski, W. Fias, A. Mousa, D. Ansari, Common and distinct brain regions  
1032 in both parietal and frontal cortex support symbolic and nonsymbolic number  
1033 processing in humans: A functional neuroimaging meta-analysis, *Neuroimage*. 146  
1034 (2017) 376–394.
- 1035 [71] A. Obersteiner, W. Van Dooren, J. Van Hoof, L. Verschaffel, The natural number bias  
1036 and magnitude representation in fraction comparison by expert mathematicians, *Learn.*  
1037 *Instr.* 28 (2013) 64–72.
- 1038 [72] Y. Ni, Y.-D. Zhou, Teaching and learning fraction and rational numbers: The origins  
1039 and implications of whole number bias, *Educ. Psychol.* 40 (2005) 27–52.
- 1040 [73] K. Kucian, U. Grond, S. Rotzer, B. Henzi, C. Schönmann, F. Plangger, M. Gälli, E.  
1041 Martin, M. von ASTER, Mental number line training in children with developmental  
1042 dyscalculia, *Neuroimage*. 57 (2011) 782–795.
- 1043 [74] B. De Smedt, L. Verschaffel, P. Ghesquière, The predictive value of numerical  
1044 magnitude comparison for individual differences in mathematics achievement, *J. Exp.*

- 1045 Child Psychol. 103 (2009) 469–479.
- 1046 [75] I.D. Holloway, D. Ansari, Mapping numerical magnitudes onto symbols: The  
1047 numerical distance effect and individual differences in children’s mathematics  
1048 achievement, *J. Exp. Child Psychol.* 103 (2009) 17–29.
- 1049 [76] L. Rousselle, M.-P. Noël, Basic numerical skills in children with mathematics learning  
1050 disabilities: A comparison of symbolic vs non-symbolic number magnitude processing,  
1051 *Cognition.* 102 (2007) 361–395.
- 1052 [77] NMAP, Foundations for success: The final report of the National Mathematics  
1053 Advisory Panel, US Department of Education, 2008.
- 1054 [78] L. Fazio, R.S. Siegler, Teaching fractions, International Academy of Education, 2011.
- 1055 [79] R. Siegler, T. Carpenter, F. Fennell, D. Geary, J. Lewis, Y. Okamoto, L. Thompson, J.  
1056 Wray, Developing Effective Fractions Instruction for Kindergarten through 8th Grade.  
1057 IES Practice Guide. NCEE 2010-4039., What Work. Clear. (2010).
- 1058 [80] S.M. Bieck, M. Splittgerber, K. Keil, R. Salvador, K. Moeller, V. Moliadze, P182  
1059 Bilateral tDCS indicates developmental differences in fraction magnitude processing,  
1060 *Clin. Neurophysiol.* 131 (2020) e117.
- 1061 [81] J.W. Stigler, K.B. Givvin, B.J. Thompson, What community college developmental  
1062 mathematics students understand about mathematics, *MathAMATYC Educ.* 1 (2010)  
1063 4–16.
- 1064 [82] L. Ma, Knowing and Teaching Elementary Mathematics: Teacher’s Understanding  
1065 OfFundamental Mathematics in China and the United States, Lawrence Erlbaum  
1066 Associates, Incorporated, 1999.
- 1067

1068

## Appendix

1069 Stimuli and corresponding parameters (i.e, fraction magnitudes, numerical distance, problem size,  
1070 matched distances and problem sizes within and between stimulus sets) for both item sets.  
1071

fraction pairs	magnitude fraction 1	magnitude fraction 2	numerical distance	problem size (PS)	matched distance within and between sets	matched PS within and between sets	Item set
$\frac{1}{8} \frac{5}{16}$	0,125	0,313	0,1875	0,438	0,022	0,078	1
$\frac{2}{3} \frac{8}{11}$	0,667	0,727	0,061	1,394	0,022	0,078	1
$\frac{3}{8} \frac{11}{13}$	0,375	0,846	0,471	1,221	0,027	0,069	1
$\frac{1}{4} \frac{22}{25}$	0,25	0,88	0,63	1,13	0,027	0,069	1
$\frac{5}{12} \frac{10}{17}$	0,417	0,588	0,172	1,005	0,006	0,066	1
$\frac{3}{14} \frac{17}{25}$	0,214	0,68	0,466	0,894	0,006	0,066	1
$\frac{3}{19} \frac{21}{29}$	0,158	0,724	0,566	0,882	0,006	0,066	1
$\frac{1}{24} \frac{16}{25}$	0,042	0,64	0,598	0,682	0,006	0,066	1
$\frac{5}{27} \frac{1}{26}$	0,185	0,039	0,147	0,224	0,008	0,045	1
$\frac{6}{23} \frac{9}{25}$	0,261	0,36	0,099	0,621	0,008	0,045	1

1072

$\frac{23}{27}$	$\frac{13}{22}$	0,852	0,591	0,261	1,443	0,016	0,062	1
$\frac{19}{23}$	$\frac{21}{25}$	0,826	0,84	0,014	1,666	0,016	0,062	1
$\frac{4}{7}$	$\frac{5}{21}$	0,571	0,238	0,333	0,810	0,013	0,087	1
$\frac{3}{7}$	$\frac{9}{22}$	0,429	0,409	0,020	0,838	0,013	0,087	1
$\frac{4}{5}$	$\frac{17}{30}$	0,8	0,567	0,233	1,367	0,022	0,017	1
$\frac{5}{6}$	$\frac{17}{22}$	0,833	0,773	0,061	1,606	0,022	0,017	1
$\frac{9}{14}$	$\frac{15}{26}$	0,643	0,577	0,066	1,220	0,00017897	0,035	1
$\frac{9}{17}$	$\frac{11}{29}$	0,529	0,379	0,150	0,909	0,00017897	0,035	1
$\frac{9}{10}$	$\frac{14}{17}$	0,9	0,824	0,077	1,724	0,00017897	0,035	1
$\frac{9}{13}$	$\frac{11}{17}$	0,692	0,647	0,045	1,339	0,00017897	0,035	1
$\frac{6}{13}$	$\frac{7}{19}$	0,462	0,368	0,093	0,830	0,006	0,033	1
$\frac{6}{11}$	$\frac{9}{26}$	0,545	0,346	0,199	0,892	0,006	0,033	1
$\frac{19}{20}$	$\frac{23}{29}$	0,95	0,793	0,157	1,743	0,001	0,002	1

$\frac{19}{25} \frac{10}{11}$	0,76	0,909	0,149	1,669	0,001	0,002	1
$\frac{1}{9} \frac{6}{25}$	0,111	0,24	0,129	0,351	0,022	0,078	2
$\frac{5}{8} \frac{7}{10}$	0,625	0,7	0,075	1,325	0,022	0,078	2
$\frac{4}{9} \frac{25}{28}$	0,444	0,893	0,448	1,337	0,027	0,069	2
$\frac{2}{9} \frac{13}{14}$	0,222	0,929	0,706	1,151	0,027	0,069	2
$\frac{1}{18} \frac{25}{26}$	0,056	0,962	0,906	1,017	0,006	0,066	2
$\frac{1}{12} \frac{10}{29}$	0,083	0,345	0,262	0,428	0,006	0,066	2
$\frac{5}{19} \frac{17}{20}$	0,263	0,85	0,587	1,113	0,006	0,066	2
$\frac{9}{20} \frac{13}{27}$	0,45	0,482	0,032	0,932	0,006	0,066	2
$\frac{5}{23} \frac{7}{30}$	0,217	0,233	0,016	0,451	0,008	0,045	2
$\frac{7}{27} \frac{1}{22}$	0,259	0,046	0,214	0,305	0,008	0,045	2
$\frac{20}{27} \frac{11}{20}$	0,741	0,55	0,191	1,291	0,016	0,062	2
$\frac{19}{21} \frac{15}{19}$	0,905	0,790	0,115	1,694	0,016	0,062	2



$\frac{2}{5}$	$\frac{3}{29}$	0,4	0,104	0,297	0,504	0,013	0,087	2
$\frac{1}{2}$	$\frac{8}{17}$	0,5	0,471	0,029	0,971	0,013	0,087	2
$\frac{3}{4}$	$\frac{13}{24}$	0,75	0,542	0,208	1,292	0,022	0,017	2
$\frac{8}{9}$	$\frac{22}{29}$	0,889	0,759	0,130	1,648	0,022	0,017	2
$\frac{7}{13}$	$\frac{11}{21}$	0,538	0,524	0,015	1,062	0,00017897	0,035	2
$\frac{9}{11}$	$\frac{21}{26}$	0,818	0,808	0,011	1,626	0,00017897	0,035	2
$\frac{8}{15}$	$\frac{13}{29}$	0,533	0,448	0,085	0,982	0,00017897	0,035	2
$\frac{9}{19}$	$\frac{11}{25}$	0,474	0,44	0,034	0,914	0,00017897	0,035	2
$\frac{8}{13}$	$\frac{9}{23}$	0,615	0,391	0,224	1,007	0,006	0,033	2
$\frac{6}{17}$	$\frac{8}{27}$	0,353	0,296	0,057	0,649	0,006	0,033	2
$\frac{20}{23}$	$\frac{12}{13}$	0,870	0,923	0,054	1,793	0,001	0,002	2
$\frac{15}{22}$	$\frac{14}{15}$	0,682	0,933	0,252	1,615	0,001	0,002	2

1075  
1076  
1077