Title
Confidence Judgments and Eye Fixations Reveal Adults‚Äô Fractions Knowledge

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 37(0)

Authors
Wall, Jenna
Thompson, Clarissa A
Morris, Bradley J

Publication Date
2015

Peer reviewed

Confidence Judgments and Eye Fixations Reveal Adults' Fractions Knowledge

Jenna Wall (jwall4@kent.edu)
Department of Psychological Sciences, P.O. Box 5190
Kent, OH 44242 USA

Clarissa A. Thompson (cthomp77@kent.edu)
Department of Psychological Sciences, P.O. Box 5190
Kent, OH 44242 USA

Bradley J. Morris (bmorri20@kent.edu)
Department of Educational Psychology, P.O. Box 5190
Kent, OH 44242 USA

Abstract

Fractions knowledge is essential to everyday life, yet many children and adults struggle to accurately represent fractions. This is the first study to investigate adults' confidence judgments and eye fixations as they solved fractions number line estimation, magnitude comparison, and magnitude ordering tasks. Educational implications are discussed.

Keywords: fractions, number line estimation, magnitude comparison, magnitude ordering, eye tracking, confidence judgments

Children's and Adults' Difficulty Representing Fractions Magnitudes

Fluency with rational numbers—fractions, decimals, and percentages—is important in the everyday lives of adults. For example, rational numbers are essential to knowing how to: double a recipe, determine the total interest paid on a mortgage, calculate the final cost of an item on sale for 75% off the original price, assess the likelihood of contracting a communicable disease, etc. Unfortunately, children and adults often struggle to accurately represent fractions (see Siegler, Fazio, Bailey, & Zhou, 2013 for a review).

What makes it so difficult to correctly represent fractions? Students often mistakenly extend their knowledge about whole numbers to fractions (Ni & Zhou, 2005; Siegler, Thompson, & Schneider, 2011). What is true of whole numbers is not true of all numbers in general. For example, 12 is larger than 9, but 1/12 is not larger than 1/9. The uniting factor, according to Siegler and colleagues' integrated theory of whole numbers and fractions, is that the magnitudes of whole numbers and fractions can be represented on number lines.

The Common Core State Standards recommend that fractions instruction begin in third grade with students representing fractions on number lines (http://www.corestandards.org/Math/Content/3/NF/). Even though fractions instruction begins early in elementary school, students continue to struggle well into adulthood to accurately represent fractions concepts—likely because it is quite difficult to inhibit the plethora of whole number knowledge that they have amassed. In one famous example (Carpenter et al., 1981), more eighth graders chose incorrect answers (19 or 21) to a simple fractions addition problem (12/13 + 7/8) than chose the correct answer (2). Only about half of sixth and eighth graders were able to correctly order fractions from smallest to largest magnitude (Mazzocco & Devlin, 2008; NCTM, 2007). Fifth graders often make errors when comparing decimal fractions (Rittle-Johnson, Siegler, & Alibali, 2001). Though it is true that all three-digit whole numbers are larger than all two-digit whole numbers, a three-digit decimal is not necessarily larger than a two-digit decimal (e.g., .539 < .68). Fifth, sixth, and even eighth graders often use non-optimal strategies as they estimate fractions on number lines. For instance, some students verbally report attending only to the denominator when estimating the location of a fraction (e.g., a student marks 3/19 closer to the 1 than to the 0 on a 0-1 number line because he noted that 19 is a large number; Siegler & Thompson, 2014; Siegler et al., 2011).

Problems representing fractions persist into high school and college. Eleventh graders failed to accurately translate between equivalent rational numbers (e.g., .029 = 29/1000, Kloosterman, 2010), and college students failed to accurately estimate the location of common numerator problems on number lines (e.g., 1/60 was placed closer to 1/1 than to 1/1440; Opfer & DeVries, 2008).

Fractions are integral to success in algebra, and success in algebra is related to access to higher education, graduation from college, and later earning capacity (National Mathematics Advisory Panel, 2008). A recent longitudinal analysis indicated that early fractions and division knowledge predicted success in algebra and overall mathematics achievement five or six years later, even after controlling for other types of mathematical knowledge, general intelligence, working memory, and family income and level of education (Siegler, et al., 2012). Accurate knowledge of fractions is crucial to later success in life.

Confidence Judgments

Confidence judgments are subjective evaluations of whether one has given a correct response to a specific problem (Dunlosky & Metcalfe, 2008). In general, judgments tend to be overconfident (i.e., judgments are higher than actual performance) when evaluating performance across a variety of cognitive tasks (Shepperd, Klein, Waters, & Weinstein, 2013). Confidence judgments are strongly influenced by the difficulty of the material being assessed in that more difficult material is associated with underconfidence (i.e., judgments lower than actual performance), while less difficult material is associated with overconfidence (Dunlosky & Metcalfe, 2008).

Confidence judgments play an important role in self-regulated learning in that confidence judgments influence the likelihood of correcting errors (Dunlosky & Rawson, 2012). Increasing the accuracy of confidence judgments appears to be an effective non-cognitive factor in improving school performance (e.g., Stankov, Morony, & Lee, 2014). Confidence judgments yield unique benefits to learning mathematics. Adults with greater numeracy and more accurate approximate number sense (ANS) were more accurate in their confidence judgments than those with lower numeracy and less accurate ANS (Winman, Juslin, Lindskog, Nilsson, & Kerimi, 2014). Children with more accurate confidence judgments achieved greater gains in mathematics than those with less accurate confidence judgments (Rinne & Mazzocco, 2014).

Eye Tracking

Researchers have used the eye-tracking paradigm to investigate children and adults' whole number knowledge. Schneider and colleagues (Schneider et al., 2008) investigated first through third graders' eye fixations as they estimated the location of whole numbers in the 0-100 range. The children fixated their gaze at the endpoints (0 and 100) and the midpoint (50) of the number line. These eye-tracking results corroborate with verbal reports that children subjectively impose reference points on number lines as they estimate the location of fractions (Siegler & Thompson, 2014; Siegler et al., 2011) and results that response times are faster when the to-be-estimated number's location is closer to subjective landmarks (Ashcraft & Moore, 2012).

There may be a mismatch between children's explicit understanding, as measured by behavioral responses and verbal reports, and implicit understanding, as measured by eye fixations (Heine et al., 2010). Though behavioral data indicated that first graders possessed a less accurate representation of numbers, their eye-tracking results were consistent with a more accurate representation.

Adults' eye movements on number line estimation tasks in the 0-1000 range were highly related to the correct location of the to-be-estimated numbers (Sullivan, Juhasz, Slattery, & Barth, 2011). These adults showed a preference for fixating near the midpoint (500) as compared to the regions around 250 and 750. Huber, Moller, and Nuerk (2014) reported eye-tracking evidence from adults as they compared fraction magnitudes and found that denominators were fixated upon more frequently than numerators.

The Current Study

The current study served as a first step to understanding the difference between adults' and children's understanding of fractions. College-aged adults completed three tasks: 1) position-to-number number line estimation, 2) magnitude comparison, and 3) magnitude ordering. After answering each problem, participants made a confidence judgment on a four-point scale ranging from not so sure to totally sure. Their eye fixations were tracked.

Hypotheses. We expected that, on average, adults would be fairly accurate on our battery of tasks, but confidence judgments and eye fixations would serve as indicators of individual differences in performance. Our study investigated three main hypotheses. First, we expected to find an association between confidence judgments and overall performance, such that participants would report feeling more confident when their estimates, comparisons, and rank ordering of fractions were more accurate. Second, both confidence judgments and eye fixations were expected to vary by problem difficulty. For instance, we anticipated lower confidence judgments and longer eye fixations for more difficult problems. "Difficult" problems were assumed to be trials in which the participant is enticed to employ a heuristic that would lead to an incorrect answer. For the magnitude comparison and ordering tasks, a difficult problem may be one in which the larger fraction has both a smaller numerator and denominator (e.g., 3/4 vs. 5/16). Participants may decide that the larger fraction is not the correct response given that its component parts are smaller in value compared to the other fraction(s). Similarly, when the larger fraction has a larger denominator (e.g., 13/17 vs. 11/15), participants may decide that the larger fraction is not the correct response based on the heuristic that all things being equal, large denominators indicate smaller fractional values. Another type of difficult problem may be one in which the fractions' decimal equivalents are close in value. Participants are more accurate and respond quicker when the magnitudes are more distant (e.g., 1/9 vs. ½ = .11 vs. .50) as compared to closer (e.g., 5/6 vs. 7/8 = .83 vs. .88). This provides evidence for the distance effect in fractions (Dehaene, Dehaene-Lambertz, & Cohen, 1998; Siegler et al., 2011). Finally, for the number line estimation task, difficult problems were considered to be trials in which the hatch mark was located far away from an experimenter-imposed landmark (0, 1) or a participant-imposed landmark (midpoint). Third, we hypothesized that adults' fixations would suggest the types of strategies used during each task. For example, we expected the number of fixations on denominators to predict performance when comparing or ordering fractions with common numerators but not fractions with common denominators. Finally, for the number line estimation task, we expected participants would fixate more frequently on the hatch mark indicated on the number line when it was located further from a subjective (e.g., midpoint) or objective (e.g., endpoint) landmark (Siegler et al, (2011) and Schneider et al. (2008).

What is the Value Added by Confidence Judgments and Eye Tracking Paradigms? To our knowledge, no previous studies have assessed adults' level of confidence as they completed fraction number line estimation, magnitude comparison, and magnitude ordering tasks in an eye-tracking paradigm. Eye tracking and confidence judgments will provide converging supporting evidence for previously reported fractions results. For instance, analyses of children's verbal reports (Siegler et al., 2011; Siegler & Thompson, 2014) have provided insights into the types of strategies that participants use to solve fractions problems, but sometimes participants find it difficult to explicitly express their thought processes. Implicit fractions understanding may outpace explicit performance (and verbal reports), and eye tracking and confidence judgments could provide some additional insights on the developmental progressions in fractions learning across the lifespan. Knowing where adults fixate when completing fractions magnitude tasks could indicate why they make the types of mistakes that they do. These insights could inform educational lessons.

Method

Participants

Fourteen undergraduate students were recruited from an introductory psychology course at a large Midwestern university (*M* age = 20 years, *SD* = 1.58; range = 18-23 years; 6 males; 78% Caucasian, 14% Asian, 7% African American). Participants received course credit.

All participants completed the magnitude comparison task before the ordering task. However, some participants (N = 8) received these tasks first followed by the number line estimation task. One participant's number line estimation data was not recorded due to equipment failure.

Procedure

Participants were seated in front of a Tobii T-60XL eye-tracker monitor. Participants were told that they would complete three tasks assessing their understanding of fractions and that their eye movements would be recorded throughout the study. They were also told that, after each trial, they would be asked to rate how confident they were in their performance. Confidence judgments were based on a 4-point scale (1-*not that sure, 2-kind of sure, 3-pretty sure, 4-totally sure*) and were reported either verbally (magnitude comparison, ordering) or electronically (number line estimation). Each participant completed a non-numerical eye-tracking calibration exercise before beginning.

Position-to-Number Number Line Estimation Participants estimated the position of a hatch mark on a number line. Each number line had a left endpoint labeled "0", a right end-point labeled "1", and a blue hatch mark corresponding to the location of one of the following fractions: 1/19, 1/15, 1/12, 1/10, 1/8, 1/7, 1/6, 1/5, 2/9, 1/4, 2/7, 3/10, 1/3, 3/8, 5/12, 4/9, 5/9, 3/5, 5/8, 2/3, 3/4, 4/5, 5/6, 7/8, 10/11, 13/14. Twenty-six number lines were presented one at a time, and participants were instructed to estimate the fraction located at each hatch mark. They responded by typing the fraction into a text box displayed below the number line. After each trial, participants were prompted to rate how confident they were in their answer. Both the fraction estimates and confidence judgments were made electronically through Qualtrics online survey software. Number lines were presented on the eye-tracking monitor, and areas of interest were created using Tobii Studio 3.2.

Magnitude Comparison Participants determined which of two fractions was larger. Forty fraction pairs (adopted from Fazio et al., 2014) were presented one at a time on the eye-tracking monitor. Each pair came from one of four ratio bins (determined by dividing the larger fraction by the smaller fraction): 1.15-1.28, 1.28-1.43, 1.48-1.65, and 2.46-2.71. Additionally, each bin included five types of trials (two of each); relative to the smaller fraction, the larger fraction had either 1) a larger numerator and an equal denominator (e.g., 7/10 and 6/10); 2) an equal numerator and a smaller denominator (e.g., 16/17 and 16/20); 3) a larger numerator and a larger denominator (e.g., 15/20 and 5/8); 4) a larger numerator and a smaller denominator (e.g., 13/14 and 12/16); or 5) a smaller numerator and a smaller denominator (e.g., 7/11 and 10/20).

Participants were encouraged to respond as quickly and accurately as possible. The fractions remained on the screen until the participant responded. After each trial, participants provided verbal confidence judgments.

Magnitude Ordering Participants ordered sets of fractions from smallest to largest. On each of ten trials, participants saw three fractions on the eye-tracking monitor, each outlined with a colored rectangle. The fraction on the left was outlined in red, the middle fraction in green, and the fraction on the right in blue. Participants made verbal responses by specifying the color of the fraction (rather than the fraction itself) when ordering (e.g., *"blue, red, green"*).

There were five types of trials (two of each); relative to the other two fractions, the larger fraction had either 1) an equal numerator and smaller denominator (e.g., 3/4, 3/15, and 3/6); 2) a larger numerator and an equal denominator (e.g., 8/9, 6/9, 3/9); 3) a larger numerator and a larger denominator (e.g., 13/17, 7/15, and 2/9); 4) a larger numerator and a smaller denominator (e.g., 10/15, 5/20, and 1/19); or 5) a smaller numerator and a smaller denominator (e.g., 4/6, 5/20, and 7/17). The fractions remained on the screen until the participant ordered all three fractions. After each trial, participants provided verbal confidence judgments. The experimenter recorded all responses. For both the magnitude ordering and comparison tasks, confidence judgments were recorded, and areas of interest were created in Tobii Studio 3.2.

Results

Hypothesis 1: Confidence & Accuracy

Position-to-Number Number Line Estimation The accuracy of number line estimates were measured by percent absolute error (PAE): PAE = (|Participant's Estimate – Correct Answer|)/Numerical Range. For example, if the location of a hatch mark corresponded to 3/4 on a 0-1 scale, and a participant estimated its location to be 2/5, the PAE would be 35% ([|.40-.75|]/1 * 100). Smaller PAEs indicate more accurate estimates. On average, participants were not very confident in their estimates (M = 2.33, SD = .73), and these judgments were not associated with overall PAE (M = 5.31, SD = 1.84%), r = -.07, $p >$ .05. Table 1 shows PAE and confidence judgments for each trial type; SDs are in parentheses in all Tables.

Table 1: Number line estimation performance

Location of Hatch Mark	PAE	Confidence (max = 4)
Close to "0"	11.11% (5.20%)	2.42 (.66)
Close to midpoint	4.50% (4.17%)	2.42 (.75)
Close to "1"	5.40% (4.76%)	2.35 (.65)
Between landmarks	3.96% (1.90%)	2.23 (.64)

Magnitude Comparison Confidence was high (M = 3.45, SD = .37) and associated with overall accuracy (M = 89.68%, SD = 12.71%), r = .753, p < .01. Table 2 shows mean accuracy and confidence judgments for each bin. Confidence judgments and accuracy were correlated within Bins 1 (r = .64, p = .014), 2 (r = .72, p < .01), and 4 (r = .65, p = .012). Confidence judgments were also related to accuracy on trials with equal numerators (e.g., 16/17 vs. 16/20), r = .80, p = < .01, when the larger fraction had both a larger numerator and denominator (e.g., 15/20 vs. 5/8), r = .57, p < .05, and when the larger fraction had both a smaller numerator and denominator (e.g., 7/16 vs. 8/21), r = .67, p = .01.

Table 2: Magnitude comparison performance

Fraction Magnitude Ratio	Accuracy	Confidence (max = 4)
Bin 1 (1.15-1.28)	86% (15%)	3.39 (.34)
Bin 2 (1.28-1.43)	87% (19%)	3.40 (.39)
Bin 3 (1.48-1.65)	94% (10%)	3.44 (.39)
Bin 4 (2.46-2.71)	91% (17%)	3.57 (.43)

Magnitude Ordering Confidence judgments were high (M = 3.31, SD = .45) and associated with overall accuracy (M = 80.24%, SD = 18.04), r = .73, p < .01. Confidence judgments were also related to accuracy on trials in which the fractions shared a common dominator, and the largest fraction had a larger numerator (e.g., 8/9 vs. 3/9 vs. 6/9), r = .82, p < .01, and when the largest fraction had a larger numerator and smaller denominator (e.g., 13/14 vs. 12/16 vs. 7/18), r = .64, p = .01. Table 3 shows mean accuracy and confidence judgments for each type of trial.

Table 3: Magnitude ordering performance by trial type

Largest Fraction Characteristics	Accuracy (max = 3)	Confidence (max = 4)
Larger Num/Equal Denom	93% (19%)	3.68 (.56)
Equal Num/Smaller Denom	83% (25%)	3.57 (.51)
Larger Num/Larger Denom	68% (33%)	2.82 (.64)
Larger Num/Smaller Denom	88% (21%)	3.39 (.66)
Smaller Num/Smaller Denom	62% (31%)	3.07 (.47)

Hypothesis 2: Confidence, Fixations, & Problem Difficulty

Position-to-Number Number Line Estimation As predicted, participants' confidence judgments were lowest on trials in which the hatch mark was between landmarks (see Table 2). Moreover, participants were significantly more confident when the hatch mark was close to "0" compared to when it was not near a landmark, t(13) = 2.21, p = .045. Interestingly, confidence was highest for "close to 0" trials even though participants were significantly less accurate on these trials compared to the other trial types (all ps < .05).

Overall, participants fixated on the hatch mark (M = 2.15, SD = .69; range = 1.07-3.79) more often than both the endpoints, "0" (M = 1.24, SD = 1.18, range = 0-6.57), t(24) = -3.13, p < .01, and "1" (M = 1.24, SD = 1.73; range = .29-5.36), t(24) = -2.85, p < .01. Contrary to our predictions, fixations on the hatch mark, midpoint, and endpoints did not vary by problem difficulty.

Magnitude Comparison As predicted, participants were more confident on "easier" trials. Table 1 shows that confidence was highest on trials with the largest magnitude differences (i.e., bins 3 and 4). Consistent with the distance effect, Bin 4 confidence judgments were significantly higher than trials within Bin 1, t(13) = -2.56, p = .024, and Bin 2, t(13) = -2.92, p = .012. Contrary to our predictions, participants were more confident when the larger fraction had both a smaller numerator and denominator (M = 3.50, SD = .43) than when both fractions shared a common denominator (M = 3.31, SD = .56), t(13) = 2.40, p = .03. Confidence was also higher when the fractions shared a common numerator compared to denominator (M = 3.53 vs. 3.31, respectively), t(13) = 2.58, p = .02.

Fixations also varied by problem difficulty. On "easy" trials, participants tended to fixate longer on the larger, correct fraction while the reverse was found for more difficult trials (see Table 4).

Table 4: Average fixation duration to correct and incorrect fraction

Largest Fraction Characteristics	Correct Fraction	Incorrect Fraction
Larger Num/Equal Denom	.95 (.16)	.81 (.18)
Equal Num/Smaller Denom	1.00 (.30)	.86 (.13)
Larger Num/Larger Denom	.89 (.28)	.89 (.30)
Larger Num/Smaller Denom	.90 (.16)	.95 (.15)
Smaller Num/Smaller Denom	.74 (.18)	.83 (.22)

Magnitude Ordering As predicted, participants' confidence judgments were significantly higher for "easy" trials (i.e., when the fractions shared a common numerator or denominator) compared to "difficult" trials (i.e., when the largest fraction had a larger denominator, or both a smaller numerator and denominator) (all *p*s < .05). Confidence was also significantly higher for trials in which the largest fraction had a larger numerator and smaller denominator compared to when it had both a larger numerator and denominator, $t(13) = 3.89$, $p < .01$.

Hypothesis 3: Fixations & Strategy Use

Position-to-Number Number Line Estimation As mentioned above, participants tended to fixate on the hatch mark more so than the endpoints or midpoint.

Magnitude Comparison Consistent with Huber et al.'s (2014) results, participants tended to fixate more on denominators ($M = 2.34$, $SD = .53$) than numerators ($M = 1.78$, $SD = .74$), $t(39) = -5.58$, $p < .01$. Contrary to our predictions, however, participants were no more likely to fixate on numerators when the denominators were equal ($M = 1.64$ and 2.37, respectively) or on denominators when the numerators were equal ($M = 2.27$ and 1.82, respectively), *p*s > .05.

Magnitude Ordering Participants tended to fixate more on numerators ($M = 4.13$, $SD = 1.05$) than denominators ($M = 3.44$, $SD = 1.28$), $t(29) = 3.15$, $p < .01$. Note that this is a pattern that differed from the magnitude comparison task results. Although fixations on the numerators did not vary by trial type, fixations on the denominator did. Interestingly, participants looked at the denominators more on trials in which the numerators were more informative (e.g., 6/9, 3/9, 8/9) compared to trials in which the denominators were more informative (e.g., 3/15, 3/4, 3/6), ($M = 2.26$, $SD = .60$, and $M = 2.11$, $SD = 1.72$, respectively), $t(13) = 2.92$, $p < .01$.

Discussion

Understanding fractions is difficult because prior knowledge about number, specifically knowledge about number magnitudes, often conflicts with correct fraction interpretation. For example, larger whole number integers indicate increasingly greater magnitudes (e.g., 20 is larger than 5), however, the same number presented as a denominator indicates a smaller magnitude (e.g., 1/20 is smaller than 1/5). The current study was the first to investigate attention to different features of fractions comparing objective, behavioral data (e.g., eye tracking) and self-reports (confidence judgments) across three different fraction tasks.

Consistent with our first hypothesis, adults' confidence judgments were associated with accuracy in both the fraction magnitude comparison and ordering tasks. This association was not found in the position-to-number line estimation task, however. Although the average PAE was fairly low, participants tended to report being less confident in their performance. It is possible that participants found the position-to-number version of the number line task to be particularly difficult and underestimated their performance. Moreover, accuracy on this task was not associated with accuracy on the other two tasks suggesting that this version of the number line estimation task may not tap the same kind of magnitude knowledge as the comparison and ordering tasks. For example, participants could have solved the number line task simply by choosing a denominator, segmenting the line based on that number, and then counting up the line until they reached the hatch mark. Future work will need to determine whether this version of the number line estimation actually taps fraction magnitude knowledge.

In line with our second hypothesis, confidence judgments varied by problem difficulty. "Easier" problems tended to receive higher confidence judgments. Fixation patterns also varied by task difficulty, but only for the magnitude comparison task. Participants tended to fixate longer on the larger fraction when it shared a common numerator or denominator with the comparison fraction, suggesting a link between strategy and accuracy.

Contrary to our third hypothesis, participants tended to look at uninformative problem components (e.g., greater looking at the denominators in the comparison task even though they were equal). Moreover, during number line estimation, participants tended to fixate on the hatch mark regardless of its distance from the endpoints or midpoint. These findings suggest that even adults sometimes use poor strategies when assessing fraction magnitude. Interestingly, the magnitude comparison and ordering tasks promote different fixation patterns. The latter task is inherently more difficult than the comparison task (i.e., three magnitude comparisons as compared to two). Perhaps this level of difficulty led participants to quickly default to an immature strategy (i.e., focusing only on numerators; see Siegler & Thompson, 2014 for a similar result). Future research will need to determine why two tasks that tap the same kind of knowledge promote different strategies (e.g., focusing on numerators vs. denominators).

The current study has educational implications and was a first step to understanding how to help children who are just beginning to learn about fractions. Given that even adults tend to focus on uninformative components when comparing fractions (e.g., common denominators), instructors should emphasize that looking at denominators, for example, is most informative when the numerators are equal. In addition, highlighting trials on which the student was confident, yet incorrect, may help learners more accurately calibrate their confidence level and more effectively allocate their resources for studying fractional values.

Acknowledgments

We would like to thank Hannah Herc for her assistance with data collection and coding.

References

Ashcraft, M. H., & Moore, A. M. (2012). Cognitive processes of numerical estimation in children. *Journal of Experimental Child Psychology, 111*, 246-267.

Carpenter, T. P., Corbitt, M. K., Kepner, H., Jr., Lindquist, M. M., & Reys, R. (1981). *Results from the second mathematics assessment of the National Assessment of Educational Progress*. Washington, DC: National Council of Teachers of Mathematics.

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neuroscience, 21*, 355–361.

Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. Sage Publications.

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, *22*(4), 271-280.

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

Heine, A., Thaler, V., Tamm, S., Hawelka, S., Schneider, M., Torbeyns, J., et al. (2010). What the eyes already "know": using eye movement measurement to tap into children's implicit numerical magnitude representations. Infant and Child Development, 19, 175–186.

Huber, S., Moeller, K., & Nuerk, H.-C. (2014). Adaptive processing of fractions-Evidence from eye-tracking. *Acta Psychologica, 148*, 37-48.

Kloosterman, P. (2010). Mathematics skills of 17-year-old students in the United States: 1978–2004. *Journal for Research in Mathematics Education, 41*, 20–51.

Mazzocco, M. M. M., & Devlin, K. T. (2008). Parts and holes: Gaps in rational number sense in children with vs. without mathematical learning disability. *Developmental Science, 11*, 681–691.

National Council of Teachers of Mathematics (NCTM) (2007). *Second handbook of research on mathematics teaching and learning*. Washington, DC: National Council of Teachers of Mathematics.

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

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

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

Rinne, L. F., & Mazzocco, M. M. (2014). Knowing Right From Wrong In Mental Arithmetic Judgments: Calibration Of Confidence Predicts The Development Of Accuracy. *PloS one*, *9*(7), e98663.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*, 346–362.

Schneider, M., Heine, A., Thaler, V., Torbeyns, J., De Smedt, B., Verschaffel, L., et al. (2008). A validation of eye movements as a measure of elementary school children's developing number sense. *Cognitive Development, 23*, 409–422.

Shepperd, J. A., Klein, W. M., Waters, E. A., & Weinstein, N. D. (2013). Taking stock of unrealistic optimism. *Perspectives on Psychological Science*, *8*(4), 395-411.

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

Siegler, R. S., Fazio, L. K., Bailey, D. H., & Zhou, X. (2013). Fractions: The new frontier for theories of numerical development. *Trends in Cognitive Sciences, 17*, 13-19.

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

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

Stankov, L., Morony, S., & Lee, Y. P. (2014). Confidence: the best non-cognitive predictor of academic achievement? *Educational Psychology*, *34*(1), 9-28

Sullivan, J., Juhasz, B., Slattery, T., & Barth, H. (2011). Adults' number-line estimation strategies: evidence from eye movements. *Psychonomic Bulletin & Review 18,* 557-563.

Winman, A., Juslin, P., Lindskog, M., Nilsson, H., & Kerimi, N. (2014). The role of ANS acuity and numeracy for the calibration and the coherence of subjective probability judgments. *Frontiers in psychology*, *5*, 1-15.