EXPLORING HOW TIME TRACKING AFFECTS STUDENT FACT PERFORMANCE

Kathryn M. Rich, Michigan State University
Meg S. Bates, University of Chicago

Suggested Citation: Rich, K. M., & Bates, M. S. (2019, April). Exploring how time tracking affects student fact performance. Paper presented at the annual meeting of the American Educational Research Association in Toronto, Ontario, Canada.

OBJECTIVES

The goal of this study is to explore the impact of different time restrictions on student performance on addition and subtraction facts. We explore this issue by examining the effects of various time-related modes on student accuracy and speed in an online fact practice-game. The results have implications for theory and practice on fact fluency development in mathematics.

THEORETICAL FRAMEWORK

Fluency with basic addition and subtraction facts is an important fundamental skill in mathematics education. Mature fact fluency is associated with better performance in mathematics (Baroody, Eiland, Purpura, & Reid, 2013; Geary, 2010; Geary et al., 2009), while problems with fact fluency are associated with mathematical performance deficiencies (Codding & Martin, 2016; Gersten et al., 2009). Although specific definitions of fact fluency differ, fact fluency is generally described as having two components: accuracy and speed (Baroody, 2006; Van der Ven, Segers, Takashima, & Verhoeven, 2017).

The specific role of speed in fact fluency is widely debated. This debate rages on two fronts: the role of speed in *assessing* fact fluency and the role of speed in *developing* fact fluency. On the assessment end, some researchers place more emphasis on speed than others. Some recent research has set a standard of high-level fact fluency as answering an addition or subtraction problem correctly in under 3 seconds (Baroody et al., 2013; Purpura, Baroody, Eiland, & Reid, 2016). Other research, however, has placed speed on equal footing with flexible and appropriate application of fact knowledge (Kling & Bay-Williams, 2014) or argued that fast mathematical performance should not be equated with good performance (Boaler, 2014). Furthermore, Kling & Bay-Williams (2014) outlined alternative ways to measure fact fluency that do not rely on assessments of speed. Despite these differences, researchers largely agree that some measure of efficiency is necessary to assess fact fluency.

The role of speed in developing fact fluency is far more contentious. Many contemporary initiatives, perhaps most notably in California (Henry & Brown, 2008), have endeavored to require fact fluency at younger grades, including memorization of facts. However, current research generally contests such an approach. Fact fluency is now seen as developing through a series of stages that include counting, reasoning strategies (in which students derive unknown facts from known facts, decomposition of numbers, number sense, and observed patterns) and, only thereafter, mastery or retrieval (Baroody, 2006). Researchers have demonstrated and argued that a premature emphasis on memorization of facts—specifically including the use of timed tests—can actually be detrimental to the learning of facts (Boaler, 2014; Henry & Brown, 2008; Kling & Bay-Williams, 2014). Instead, they have called for instruction that emphasizes exploration, number sense, and meaningful strategy use (Baroody, 2006). Research has specifically called out the potential of timed tests to exacerbate math anxiety (Boaler, 2006); if true, timed tests might lead to lower accuracy, as math anxiety has been shown to inhibit working memory (and thus recall) in students with high math potential (Ramirez, Gunderson, Levine, & Beilock, 2013).

Given that the end goal of fact fluency development involves accurate and efficient fact retrieval, questions remain about the role of time in the development of fluency. At some point, time must be taken into consideration in student fact learning. But under what circumstances, if any, can attention to time be useful for student development of fluency? This study explores this question by examining student performance in a fact-practice game with multiple modes—some timed and some untimed.

DATA SOURCE

Our data comes from a web-based facts game associated with *Everyday Mathematics,* a research-based Pre-Kindergarten to Grade 6 mathematics curriculum developed at the University of Chicago and published by McGraw-Hill Education (University of Chicago School Mathematics Project, 2007, 2012). The game is intended primarily for students in grades 1-4, where most facts instruction and practice appears in the curriculum materials. Students answer addition and subtraction questions presented with reference to fact families, meaning that related addition and subtraction facts are presented together. For example, students might be asked to choose a set of four equations (a fact family) that corresponds to a domino. A domino with 4 dots on one side and two on the other would be associated with the fact family 4 + 2 = 6, 2 + 4 = 6, 6 – 4 = 2, and 6 – 2 = 4. In each round, students answer questions until they answer 10 correctly. Importantly for this study, students choose a mode in which to play the game, with each mode associated with a form of timing.

We received the de-identified data, collected between August 2011 and September 2016, from representatives at McGraw-Hill Education in the fall of 2016. The data was structured around individual question attempts. Each question presented to a student was recorded as a separate row in the data table along with an identification number for the student, the mode in which the game was played, the time in milliseconds spent on the question, and whether the answer given was correct or incorrect. For the purposes of our analysis, we separated the data by mode, then aggregated by student.

METHODS

Independent and Dependent Variable Selection. The independent variable was the timing mode in which the fact problems were answered. Table 1 describes three timing modes that are available in the game. Henceforce we use the Time Codes in Table 1 to refer to the timing modes.

Table 1. Summary of Timing Modes

Time Code	Role of Time Tracking in Mode
No Time	No time limit, and time to answer does not factor into game score.
Beat Your Time	No time limit, but total time to complete a game is compared to time to complete previous games.
Time Limit	Time limit in place; failure to answer within the time limit (6 seconds) is treated as an incorrect response.

The dependent variables were *accuracy* and *speed. Accuracy* is defined as the number of questions a student attempted in each timing mode divided by the number of those questions that the student answered correctly. An accuracy measure of 1 therefore represents perfect performance, with higher measures indicating lower accuracy. *Speed* is defined as the average number of seconds a student took to answer a question *correctly* in a given timing mode. Questions that were answered incorrectly were excluded from the speed analysis; we did not judge time to answer a question incorrectly as a meaningful measure of student performance.

Sample. The original data set contained data from more than 150,000 students. The game was available to all *Everyday Mathematics* users, and so students may have been in any grade level K-6. Specific information about the grade level of each student was unavailable, but we expect that the addition and subtraction game would be used by teachers of grades 1-4. To enable paired comparison of student performance across mode, we included only students who completed at least one round of the game in each of the three timing modes. This included 886 students.

Analysis Plan. We conducted statistical tests to detect overall possible effects of timing mode on students' accuracy and speed, followed by post-hoc pairwise comparisons of modes. Examination of the data showed that the scores were skewed and overdispersed, thereby violating the normality assumptions of a repeated-measures ANOVA and dependent samples t-test. As such, we conducted non-parametric alternatives to theses tests: Friedman Rank-Sum tests and pairwise Wilcoxon Signed Rank tests, respectively. For the Wilcoxon tests, we used a Bonferroni adjustment. To estimate effect size, we calculated Cliff's delta, which is a non-parametric alternative to Cohen's d.

RESULTS

Accuracy. The mean and median accuracy across students in each timing mode are shown in Table 2. We report both the mean and median due to our use of non-parametric tests. A Friedman Rank Sum test detected an overall difference across groups ($p < 0.001$, chi-squared = 497.54, df = 2). Follow-up pairwise Wilcoxon tests showed significant differences in all comparisons ($p = 0.002$ for No Time versus Beat Your Time, $p < 0.001$ for other comparisons). Calculation of Cliff's delta suggested a negligible practical effect between No Time and Beat Your Time (delta estimate: 0.05). Cliff's delta indicated a medium practical effect between the Time Limit mode and both other modes (delta estimates of 0.42 and 0.47). In short, a time limit is associated with lower accuracy than both the No Time and Beat Your Time modes, but there was no practical difference between the No Time and Beat Your Time modes.

Table 2: Mean and Median Accuracy by Timing Mode

Time Code	Mean Accuracy	Median Accuracy	SD
No Time	1.23	1.11	0.30
Beat Your Time	1.25	1.13	0.33
Time Limit	1.48	1.33	0.47

Speed. The mean and median speed (in seconds) to produce a correct answer according to timing mode are shown in Table 3. A Friedman Rank Sum test detected an overall difference across groups ($p < 0.001$, chi-squared = 1041.6, df = 2). Follow-up pairwise Wilcoxon tests showed significant differences ($p < 0.001$) in all comparisons. Calculation of Cliff's delta showed a medium practical effect in comparing Beat Your Time and No Time (delta estimate: 0.35). The remaining delta estimates indicate large effect of Timed mode over both other modes (delta estimates of 0.59 and 0.82). In summary, the Timed mode is associated with the fastest student response times, with a large effect over all other modes. The Beat Your Time mode is associated with the next-fastest response times, with a medium effect over the No Time mode.

Table 3. Mean and Median Speed by Timing Mode

Time Code	Mean Speed (s)	Median Speed (s)	SD
Timed	3.68	3.49	1.10
Beat Your Time	6.42	5.21	5.32
No Time	9.02	7.23	9.25

SCHOLARLY SIGNIFICANCE

This study contributes to the existing literature debating the effects of timed drills in student fact learning. Specifically, our analyses showed that the Timed mode was associated with the lowest overall rate of accuracy, with a medium practical effect differentiating it from all other modes. The Timed mode was also associated with the fastest average time to produce a correct answer, with a large practical effect over the other modes.

This result is consistent with the theory that time limits are detrimental to fact learning (Boaler, 2014; Henry & Brown, 2008; Kling & Bay-Williams, 2014). Although students appear to produce correct answers more quickly with a time limit in place, they also answer significantly more facts incorrectly. However, a number of features of our data set lead us to interpret these results with caution. First, it should be noted that a unique feature of the Timed mode is that it records non-responses from students as incorrect answers (when the time limit expires), whereas the other modes will not record non-responses but rather wait for a student response before scoring. A portion of the incorrect answers, then, were actually non-responses that are not possible in other modes. This may have contributed to the overall lower rate of accuracy. With respect to speed, because any answer taking longer that 6 seconds for a student to produce would be marked as incorrect in this mode, the times to produce an incorrect answer are necessarily restricted to a range of 0 to 6 seconds. This certainly contributed to an overall faster average time in producing correct answers.

Even so, prior research suggests that time limits are detrimental to fact learning (Boaler, 2014; Henry & Brown, 2008; Kling & Bay-Williams, 2014), and our results are not inconsistent with this theory. Interestingly, our results further suggest that it is not simply the awareness of time tracking that leads to a detriment in accuracy. There were no practical differences in the accuracy scores for the Beat Your Time mode compared the No Time mode. Thus, despite being aware that their time to answer was being recorded, students performed equally well in the Beat Your Time mode as when they believed their response times were not being tracked. This suggests that awareness of a time *limit*, as opposed to awareness of time *tracking*, could be a contributing factor to poorer student fact performance. This may indicate that time limits, in particular, cause math anxiety that interferes with working memory (see Ramirez et al., 2013).

With respect to speed, Beat Your Time mode was associated with faster times to produce a correct answer than the No Time mode. Taken together with the accuracy results for the Beat Your Time mode, this suggests the following important finding: Tracking of time, without imposing a time limit, was associated with faster accurate responses from students, with no detriment to overall accuracy. This suggests that time tracking, where students are challenged to beat their own previous speed records as opposed to an externally set time limit, may be a productive instructional strategy for developing fact fluency.

REFERENCES

Baroody, A. J. (2006). Why children have difficulties mastering the basic number combinations and how to help them. *Teaching Children Mathematics*, *13*(1), 22-31.

Baroody, A. J., Eiland, M. D., Purpura, D. J., & Reid, E. E. (2013). Can computer-assisted discovery learning foster first graders' fluency with the most basic addition combinations? *American Educational Research Journal, 50*(3), 533-573.

Boaler, J. (2014). Research suggests that timed tests cause math anxiety. *Teaching Children Mathematics, 20*(8), 469-474.

Codding, R. S., & Martin, R. (2016). Tier 3: Intensive mathematics intervention strategies. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of Response to Intervention* (2nd ed.), pp. 375-388. New York: Springer.

Geary, D. C. (2010). Mathematical disabilities: Reflections on cognitive, neuropsychological, and genetic components. *Learning and Individual Differences, 20,* 130-133.

Geary, D. C., Bailey, D. H., Littlefield, A., Wood, P., Hoard, M. K., & Nugent, L. (2009). First-grade predictors of mathematical learning disability: A latent class trajectory analysis. *Cognitive Development, 24*(4), 411-429.

Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). *Assisting students struggling with mathematics: Response to Intervention (RtI) for elementary and middle schools* (NCEE 2009-4060). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies. ed.gov/ncee/wwc/publications/practiceguides/.

Henry, V., & Brown, R. (2008). First-grade basic facts: An investigation into teaching and learning of an accelerated, high-demand memorization standard. *Journal for Research in Mathematics Education, 39*(2), 153-183.

Kling, G., & Bay-Williams, J. M. (2014). Assessing basic fact fluency. *Teaching Children Mathematics, 20*(8), 488-497.

Purpura, D. J., Baroody, A. J., Eiland, M. D., & Reid, E. E. (2016). Fostering first graders' reasoning strategies with basic sums: The value of guided instruction. *The Elementary School Journal, 117*(1), 72-100.

Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2013). Math anxiety, working memory, and math achievement in early elementary school. *Journal of Cognition and Development, 14*(2), 187-202.

University of Chicago School Mathematics Project. (2007). *Everyday Mathematics 3.* New York: McGraw-Hill Education.

University of Chicago School Mathematics Project. (2012). *Everyday Mathematics Common Core State Standards Edition.* New York: McGraw-Hill Education.

Van der Ven, F., Segers, E., Takashima, A., & Verhoeven, L. (2017). Effects of a tablet game intervention on simple addition and subtraction fluency in first graders. *Computers in Human Behavior, 72,* 200-207.