ReasoningMind

Reasoning Mind
DISD 2015 Report

Victor Kostyuk, Leigh Mingle,
Steven Gaudino, Diedre Douglas

Solving for every variable

RR-1409

Contents

REASONING MIND DISD 2015 REPORT

Executive Summary

- In 2014–15, Reasoning Mind was used as a supplementary mathematics curriculum districtwide in grades 2–3, by about 70% of schools in the district in grade 4, and as a pilot core curriculum by several schools in grade 5.
- DISD implementation **targets of time online[1] were met by fewer than half the students** who used Reasoning Mind and **accuracy targets[2] were met by just over half the students.**
- Compared to 2013–14, there was a **small improvement in the number of students meeting targets of time online** in grades 2–4, and a **moderate improvement of students meeting the accuracy target in grade 2.**[3]
- District surveys showed that teachers had a favorable view of Reasoning Mind's effects on students, with **86 percent of supported teachers and 71 percent of nonsupported teachers agreeing that students benefit from Reasoning Mind.** Campus administrators had a less favorable view, with 58 percent agreeing Reasoning Mind benefits students.[3]
- The **level of positive opinion of Reasoning Mind was higher among supported teachers** than among non-supported teachers across all survey questions.[3]
- Controlling for school differences, in-system performance, SES, and previous Iowa/STAAR performance, **in grades 2–4**, there was a **statistically significant ($p<0.05$) positive association between time online and performance on the Iowa Assessment (grade 2) and 2015 STAAR (grades 3–4).** In grade 5, the association was positive and marginally statistically significant ($p<0.1$).
- Controlling for school differences, SES, and previous STAAR performance, students who **used Reasoning Mind for at least 0.1 hours** during the year (a very low bar) had an associated **increase in STAAR performance that was negligible in grade 4 and 0.9 questions in grade 5 (neither statistically significant).**
- Controlling for school differences, SES, and previous STAAR performance, students who **used Reasoning Mind for at least 75% of the DISD's target hours online** had an associated **increase in STAAR performance of 0.85 and 1.4 questions in grades 4 and 5, respectively (both statistically significant).**

Overall, the results of the 2014–2015 school implementation of Reasoning Mind in Dallas ISD indicate that the use of the program is associated with better student outcomes as measured by the Iowa Assessment (grade 2) and STAAR (grades 3–5). The positive effects have been registered in all grades 2–5 and as the use of the program approaches target time online, the increase in STAAR performance becomes statistically and educationally significant.

[1] 60 hours in grades 2–4 and 205.33 hours in grade 5
[2] 75% accuracy on basic-level problems (level A problems in Guided Study)
[3] See Bush and Kim (2015)

1. Introduction and Background

Dallas ISD first partnered with Reasoning Mind in the 2009–2010 school year, enrolling 2,321 grade 3 and grade 4 students. In 2010–2011, enrollment at the district expanded to 3,352 students; seeing strong results for these students, the district made the bold decision to implement Reasoning Mind in 2011–2012 with all of the district's grade 2 students. In 2012–2013 this was expanded to all of the district's grade 2 and grade 3 students and in 2013–2014 schools were given the choice to use Reasoning mind with their grade 4 students. During the 2013–2014 and 2014–2015 school years, over 70 percent of schools in Dallas ISD opted to use Reasoning Mind for their grade 4 students. In 2014–2015, eleven schools[4] were initially assigned to use Reasoning Mind as the core curriculum for their grade 5 students. Six of these schools used Reasoning Mind throughout the whole year. This collection of schools was intended as a pilot of the core 5th-grade curriculum in the district.

In the 2014–2015 school year, teachers had positive attitudes towards Reasoning Mind. Surveys of Dallas ISD teachers conducted by Bush and Kim (2015) showed approval for using Reasoning Mind, with 80% of supported and 62% of non-supported teachers indicating they would like to use the program again next year. Additionally, 86% of supported and 71% of non-supported teachers said Reasoning Mind benefits students, 70% of supported 51% of non-supported teachers said Reasoning Mind helps them be more effective, and 80% and 62% of supported and non-supported teachers said they would recommend Reasoning Mind to others.

Campus administrators were less supportive of the program than teachers in the 2014–2015 school year: 47% of administrators said they would like to use Reasoning Mind again next year, 58% said it benefits students, 39% said it helps teachers be more effective, and 43% said they would recommend Reasoning Mind to others. In the remainder of this report, we investigate the association between Reasoning Mind and student performance on the State of Texas Assessment of Academic Readiness (STAAR). We use several statistical methods to provide an estimate of the benefits of the program.

All of the following analyses are correlational. Because use of Reasoning Mind was either the district or school's decision, it was impossible to randomly assign schools or students to use the program. Without random assignment, we cannot be sure that all possible confounding factors are accounted for (Shadish, Cook, & Campbell, 2002). To mitigate the impact of confounding factors, we statistically controlled for key available variables; however, there is no way to ensure that there are no other factors influencing student performance. As such, we cannot make any claims about a causal relationship between Reasoning Mind and student performance.

[4] Adams, Henderson, Kleberg, Medrano, Starks, Stemmons, Cochran, Salazar, Tolbert, Roberts, and Kennedy.

2. Performance and the use of Reasoning Mind

Given that time in Reasoning Mind takes the place of time spent in whole class instruction[5], we would like to see whether time in Reasoning Mind is positively correlated with improved performance. Additionally, in grades 4 and 5, many schools in DISD did not use Reasoning Mind. Controlling for differences among the schools selected to use Reasoning Mind and those that did not, we evaluate the effect on STAAR performance associated with the use of the program.

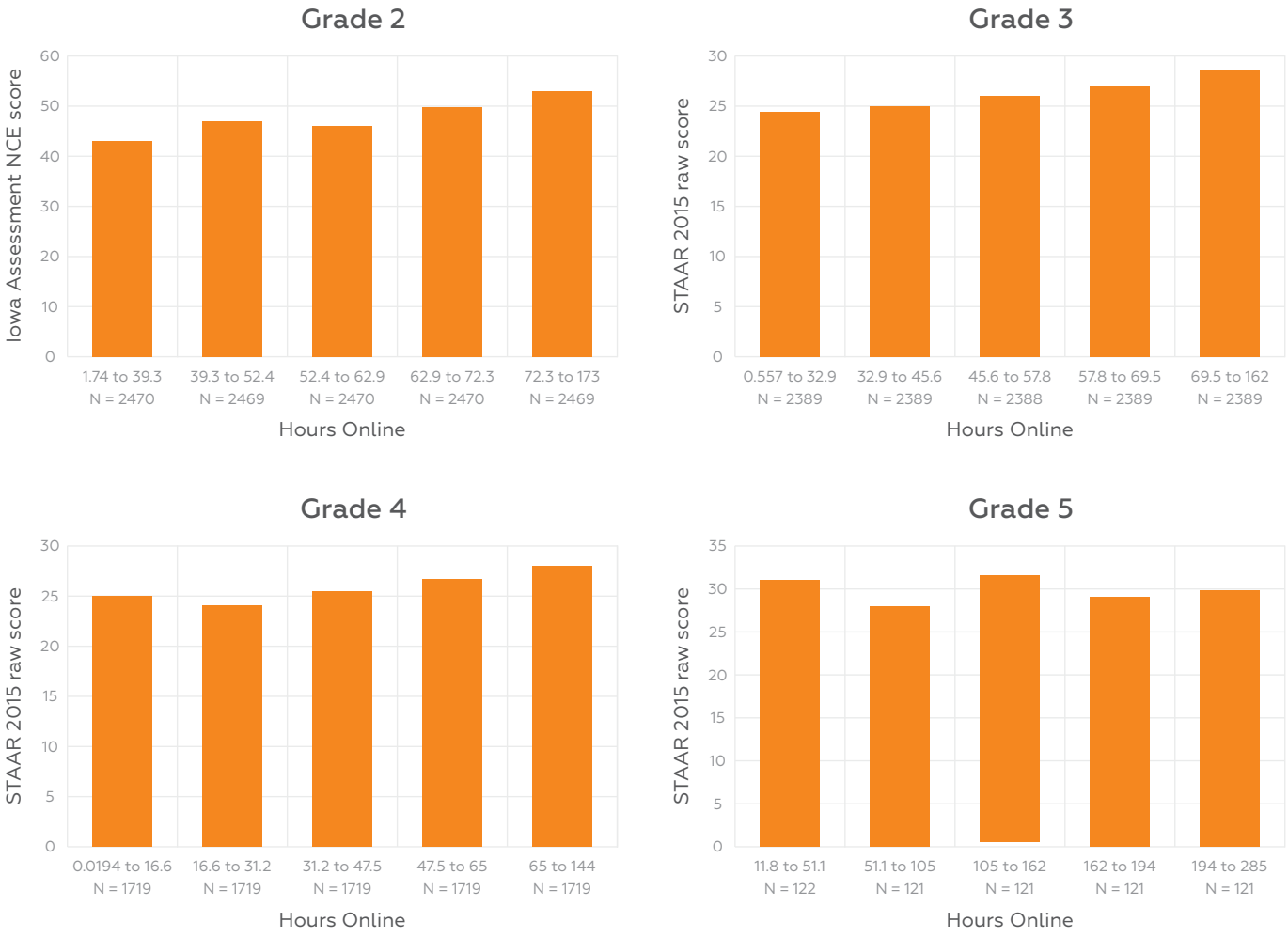
Relationship of Performance to Time Online

A key implementation target in DISD is time online in Reasoning Mind. Dallas ISD students have 90 minutes a day scheduled for mathematics, which translates into 262 hours of instruction a school year. In grades 2 through 4, Reasoning Mind is an intensive supplemental program. During the 2014–2015 school year, Dallas ISD set a goal of 60 hours of Reasoning Mind for all supplemental students. This represents 23% of instructional time for the year. According to the 2014–2105 report conducted by Dallas ISD (Bush & Kim, 2015), only 44% of grade 2 students, 35% of grade 3 students, and 26% of grade 4 students met this goal.

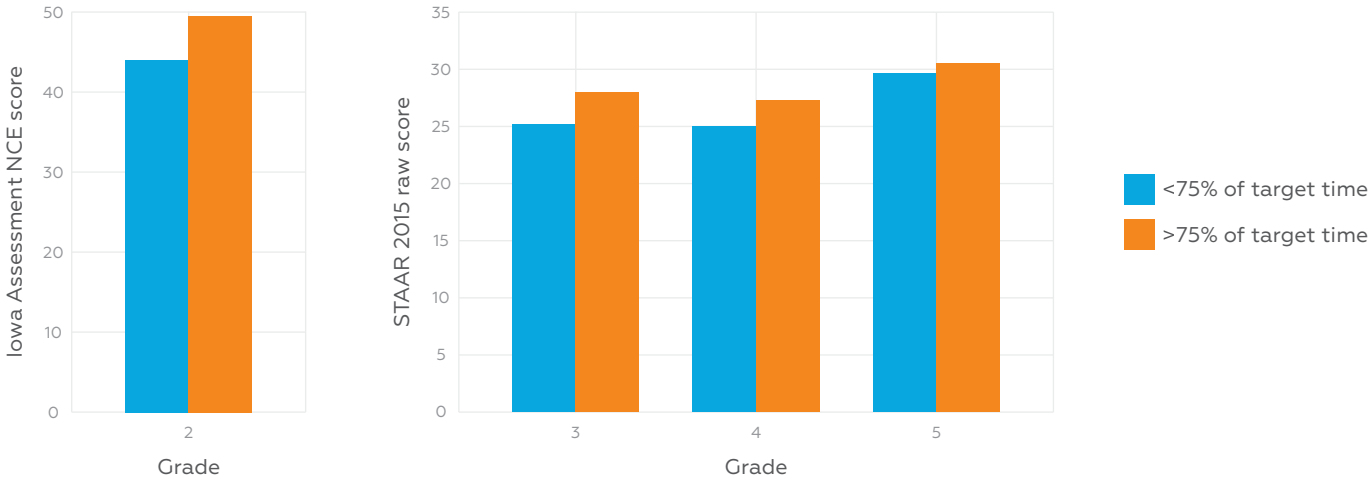
In grade 5, Reasoning Mind is designed as a core curriculum, intended to be used for at least 60 minutes per day. In accordance with this recommendation, Dallas ISD set a goal of 205 hours of Reasoning Mind for grade 5 students. This represents 78% of instructional time for the year. Only 21% of the grade 5 students in Dallas ISD met this goal.

For all following analyses, we define a **Reasoning Mind student** as a DISD student appearing in the PEIMS October 2014 snapshot who used Reasoning Mind for at least 0.1 hours during the academic year. Below, we examine the performance of all Reasoning Mind students on the Iowa Assessment (grade 2) and the 2015 STAAR (grades 3–5) according to quintiles of time online in Reasoning Mind. With the exception of grade 5, there was a clear pattern of positive correlation between time online and standardized test performance.

[5] In addition to using Reasoning Mind in school, students can and do access Reasoning Mind outside of school hours. In Appendix 1 we provide summary statistics on outside of school Reasoning Mind use. Neither the graphical pattern of correlation nor the results of the HLM analyses change in significance or sign when hours in school are used instead of total hours.



Because time online is an important metric both for Reasoning Mind and Dallas ISD, we also compare the performance of Reasoning Mind students who did not meet 75% of the target hours online (who used the system for less than 45 hours in grades 2–4 and less than 154 hours in grade 5) with those who met or exceeded 75% of the targeted hours. In all grades, Reasoning Mind students who used the system for at least 75% of the targeted hours outperformed their counterparts who used the system less.



Correlation of Performance with Time Online: HLM Analysis

In most statistical analyses, there is an assumption that the individual observations are independent of one another. When study participants are grouped together in similar environments, such as students within schools, these observations are not independent. Students within the same school behave more similarly to one another than to students in different schools (Snijders and Bosker, 1999). This dependency makes the use of traditional statistical analyses unreliable.

> "The correlated errors among the individuals within a group violate the independent observations assumption of ordinary least squares (OLS) estimation, resulting in downwardly biased standard error estimates, overly large test statistics, and inflated Type I error rates" (Krull & MacKinnon, 2001, p. 251; see also, Barcikowski, 1981; Moulton, 1986; Scariano & Davenport, 1987; Scott & Holt, 1982; Walsh, 1947).

To address this issue, a hierarchical linear model can be used to control for group effects.

In the Dallas ISD data for 2014–2015, a significant part of the variance in student performance in grades 2–5 is accounted for by school level variables. To determine how much of the variance can be attributed to school level factors, we calculated the intraclass correlation (ICC), the ratio of between group to total variance, for each grade.

Grade	ICC	F	p-value
2	0.12	F (149, 12848) = 11.98	<.001
3	0.14	F (149, 11922) = 13.09	<.001
4	0.17	F (150, 11511) = 16.26	<.001
5	0.15	F (150, 10838) = 12.58	<.001

The significance of variation at the school level indicates the need for a hierarchical linear model (HLM) to account for this variance (Snijders and Bosker, 1999).

To account for variation resulting from school level clustering, we use the following HLM[6]:

Level 1 (student): $CurrScore_{ij} = \beta_{0j} + \beta_1 Hours_{ij} + \beta_2 Acc + \beta_3 PrevScore_{ij} + \beta_4 SES_{ij} + e_{ij}$

Level 2 (school): $\beta_{0j} = \gamma_{00} + u_{0j}$

where $e_{ij} \sim N(0, \sigma^2)$ and $u_{0j} \sim N(0, \tau^2)$

- *CurrScore* is the 2015 Iowa Assessment NCE or STAAR raw score
- *PrevScore* is the 2014 Iowa Assessment NCE or STAAR raw score
- *Hours* is the total number of hours spent in Reasoning Mind
- *Acc* is accuracy on level A problems
- *SES* is a binary (true/false) variable marking economically disadvantaged status

[6] Another way to model the the nesting of students within schools is to treat school intercepts as "fixed effects" – use categorical indicator variables in a single–level generalized linear regression model. There is some debate in the literature on when to use fixed effects and when to use random effects for nesting within clusters with no cluster–level predictors. Clark and Linzer (2015) address this question in great depth, providing some criteria and recommendations. We chose to follow Gelman and Hill (2007, p. 245) on this question: "Our advice is to *always* use multilevel modeling ("random effects")." [emphasis in the original]

	Dependent variable:			
	Iowa Assessment score	STAAR 2015 score		
	Grade 2	Grade 3	Grade 4	Grade 5
FIXED EFFECTS				
Hours Online	0.12*** (0.01)	0.06*** (0.00)	0.02*** (0.00)	0.02* (0.01)
Level A accuracy	0.55*** (0.01)	0.58*** (0.01)	0.18*** (0.01)	0.29*** (0.04)
ITBS 2014 score	0.46*** (0.01)			
STAAR 2014			0.58*** (0.01)	0.65*** (0.03)
Low SES	−1.07* (0.63)	−1.94*** (0.23)	−1.17*** (0.34)	0.38 (1.24)
Constant	−15.16*** (1.15)	−21.57*** (0.58)	−4.35*** (0.59)	−16.93*** (3.01)
RANDOM EFFECTS				
Schools	148	148	124	9
School variance	24.3	7.09	5.81	7.01
Residual variance	174.5	43.33	28.06	27.78
Observations	11,358	11,732	7,413	575
Log Likelihood	−45,616.00	−38,952.00	−23,035.00	−1,789.00
Akaike Inf. Crit.	91,245.00	77,916.00	46,084.00	3,591.00
Bayesian Inf. Crit.	91,297.00	77,960.00	46,132.00	3,622.00
Note:				*p < 0.1; **p < 0.05; ***p < 0.01 [7]

These results show that hours online have a positive impact on assessment performance in all grade levels, statistically significant at the 0.05 level in grades 2–4, after accounting for school effects, previous performance, SES, and in-system level A accuracy.

[7] Note that the grade 5 model shows students in nine schools since 5th-grade students in nine schools used the Reasoning Mind program. Three of these schools only used the program in the Fall semester, and six schools used it throughout the whole year.

Reasoning Mind and non-Reasoning Mind Students in Grades 4 and 5

To assess the effect of using Reasoning Mind in grades 4 and 5, we control for school level variation, previous performance, and SES, by using the following HLM model:

Level 1 (student): $CurrScore_{ij} = \beta_{0j} + \beta_1 PrevScore_{ij} + \beta_2 SES_{ij} + \beta_3 RM_{ij} + e_{ij}$

Level 2 (school): $\beta_{0j} = \gamma_{00} + u_{0j}$

where $e_{ij} \sim N(0, \sigma^2)$ and $u_{0j} \sim N(0, \tau^2)$

- CurrScore is the 2015 STAAR raw score
- PrevScore is the 2014 STAAR raw score
- SES is a binary (true/false) variable marking economically disadvantaged status
- RM is a binary variable true when the student is a Reasoning Mind student (used at least 0.1 hours) and false otherwise, in the two left models below, or true when the student used Reasoning Mind for at least 75% of the targeted time online and false otherwise, in the two right models below[8]

The number of Reasoning Mind students and of students who used Reasoning Mind for 75% of the targeted time online or longer in grades 4 and 5 is shown in the table below.

	Total students[9]	Reasoning Mind students[10]	Reasoning Mind students exceeding 75% of time online target
Grade 4	10,845	7,984	3,468
Grade 5	10,321	576	261

[8] Note that RM is introduced as a level 1 (student level) and not a level 2 (school level) predictor. We considered using RM as a school level predictor, but since some schools had only some of their students use Reasoning Mind, we decided it would be more accurate to have Reasoning Mind status be a student level predictor.

[9] This is the total number of students in the PEIMS October 2014 snapshot who had prior and current year STAAR scores and low SES indicator data. I.e., the total number of students considered in this analysis.

[10] As defined on page 5, these are all DISD student appearing in the PEIMS October 2014 snapshot who used Reasoning Mind for at least 0.1 hours during the academic year.

	Dependent variable:			
	All Reasoning Mind students		Exceeding 75% time online target	
	Grade 4	Grade 5	Grade 4	Grade 5
FIXED EFFECTS				
STAAR 2014 score	0.736*** (0.006)	0.835*** (0.006)	0.734*** (0.006)	0.835*** (0.006)
Low SES	−1.256*** (0.268)	−0.678** (0.275)	−1.275*** (0.267)	−0.673** (0.275)
Reasoning Mind student	0.002 (0.304)	0.930 (0.760)		
More than 75% of target			0.846*** (0.196)	1.352** (0.674)
Constant	6.454*** (0.436)	4.152*** (0.394)	6.239*** (0.382)	4.167*** (0.393)
RANDOM EFFECTS				
Schools	150	151	150	151
School variance	6.52	6.29	6.54	6.24
Residual variance	31.83	35.99	31.78	35.98
Observations	10,845	10,321	10,845	10,321
Log Likelihood	−34,359.000	−33,326.000	−34,350.000	−33,325.000
Akaike Inf. Crit.	68,730.000	66,665.000	68,712.000	66,662.000
Bayesian Inf. Crit.	68,774.000	66,708.000	68,756.000	66,706.000
Note:				*p < 0.1; **p < 0.05; ***p < 0.01

Thus, after controlling for school effects, previous performance, and SES, we see that in both grades 4 and 5 using Reasoning Mind is associated with a positive but not statistically significant effect on performance. However, using Reasoning Mind for at least 75% of the target time – at least 45 hours in grade 4 and 154 hours in grade 5 – is associated with a statistically significant improvement of 0.85 and 1.35 questions in grades 4 and 5, respectively.

3. Continued Reasoning Mind Commitment to the Success of the Partnership with DISD

Based on the results in Bush and Kim (2015) and our analysis, Reasoning Mind will act to strengthen Reasoning Mind implementation at all grade levels.

Promoting the Reasoning Mind Community. The Reasoning Mind Community is an online space where teachers and administrators can see their progress, share best practices, download resources, sign up for classes, find answers to questions, propose ideas, and get immediate help through live chat. By taking advantage of this new site, Dallas ISD will have convenient access to the tools and resources they need to effectively implement Reasoning Mind.

Emphasizing performance on basic Reasoning Mind problems. Performance on basic (A-level) problems are displayed prominently in the Reasoning Mind Community dashboard. In addition, Reasoning Mind and Dallas ISD will work together to support student progress on this metric.

Emphasizing performance on advanced problems. Performance on advanced (B- and C-level) problems are also key metrics on the Reasoning Mind Community dashboard. Reasoning Mind will continue to encourage student productivity and work in the Wall of Mastery to ensure their exposure to these higher level problems.

Emphasizing time in Guided Study. Time online has been a major push from Dallas ISD's central administration in past years. Reasoning Mind will continue to work with school and district administrators to support an early launch, a quick resolution of IT issues, and a commitment to implementation fidelity.

Emphasizing productivity. Productivity is a new metric displayed on the Reasoning Mind Community. We will work closely with Dallas ISD administrators and teachers to introduce and monitor this important measure of student progress.

Adding flexibility to support. To effectively promote, monitor, and take action on the metrics above, we will use a flexible model of support where we allocate our resources in response to dashboard data rather than predetermined assumptions.

Focusing our support. We will continue to prioritize our 5th grade implementations. In addition to putting our strongest Implementation Coordinators on those campuses, we have prioritized their time to provide close and comprehensive support.

Promoting consistent testing. Requiring an assessment built around a different curriculum sequence created significant anxiety and turnover in the 14–15 academic year. Moreover, it did not serve as an accurate assessment of what students knew because it did not test what students studied most. Reasoning Mind and Dallas ISD will continue to work on a solution to the problems posed by the ACP.

Refining the implementation of STAAR Readiness. Because the foundation built in Guided Study pointed to a greater impact than STAAR Readiness, we will continue to emphasize the importance of time in Guided Study. At the same time, we will analyze the way STAAR Readiness was used in the past year and identify the best practices of those who used it with the most success. Lastly, we will use the TEKS5 curriculum in the upcoming year so that STAAR Readiness modules are used in a more intentional and coherent manner.

Investigating and improving the quality of STAAR Readiness materials. The curriculum team will review STAAR problems recently released by TEA and compare them to STAAR Readiness problems on corresponding standards. Improvements will be made to ensure that STAAR Readiness problems and instructional materials match the format and logic of the problems on the test. A preliminary analysis of the released problems has shown that for most standards, our STAAR Readiness materials closely reflect the released problems.

4. References

Barcikowski, R. S. (1981). Statistical power with group mean as the unit of analysis. *Journal of Educational Statistics, 6, 267–285.*

Bush, J. & Kim, M. (2015). Evaluation of Reasoning Mind: Final Report 2014–15.

Clark, T. S. & Linzer, D. A. (2015). Should I Use Fixed or Random Effects? *Political Science Research and Methods, 3(02), 399–408.*

Gelman, A. & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Cambridge: Cambridge University Press.

Krull, J. L., & McKinnon, D. P., (2001). Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research, 36, 249–277.*

Moulton, B. R. (1986). Random group effects and the precision of regression estimates. *Journal of Econometrics, 32, 385397.*

Scariano, S. M. & Davenport, J. M. (1987). The effects of violations of independence assumptions in oneway ANOVA. *The American Statistician, 41, 123–129.*

Scott, A. J. & Holt, D. (1982). The effect of twostage sampling on ordinary least squares methods. *Journal of the American Statistical Association, 77, 848–854.*

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasiexperimental designs for generalized causal inference.* Belmont, CA: Wadsworth.

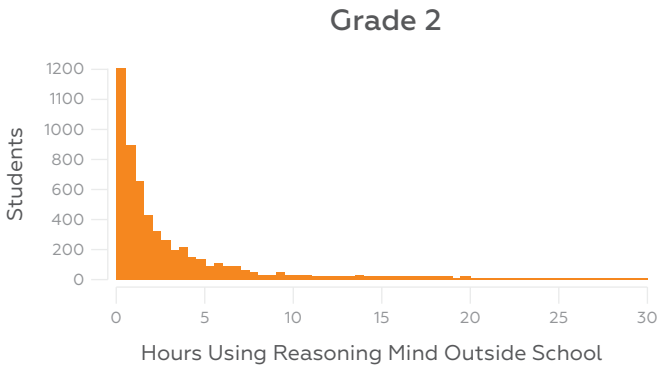
Snijders, T. & Bosker, R. (1999). *Multilevel analysis.* London: Sage.

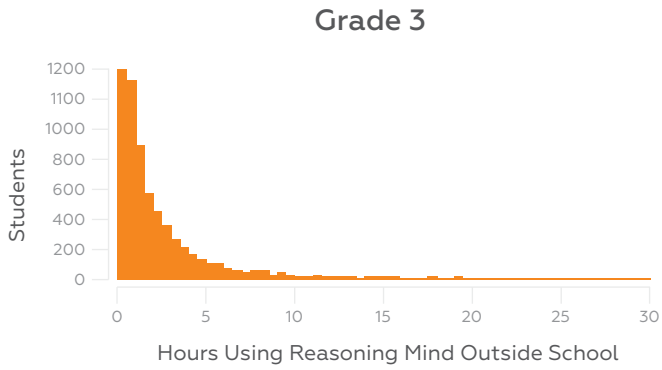
Walsh, J. E. (1947). Concerning the effect of the intraclass correlation on certain significance tests. *Annals of Mathematical Statistics, 18, 88–96.*

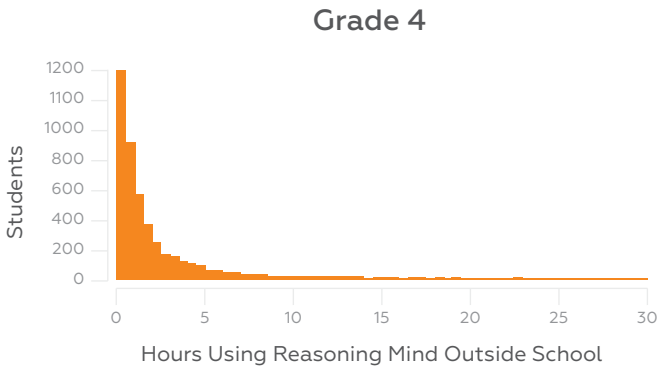
Appendix 1: Use of Reasoning Mind Outside of School Hours

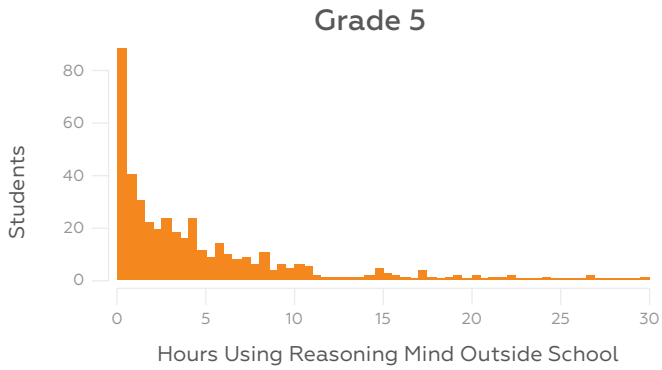
Grade	Total RM students	Used RM outside school hours	Mean hrs outside school for RM students who used	S.d.	Max
2	12348	6752	3.3	5.8	69.0
3	11944	8190	2.6	4.6	81.6
4	8575	5284	1.8	3.5	65.8
5	606	496	4.6	6.1	61.0

Distribution of outside school hours usage:









Graphs showing the relationship of time online during school hours to performance on the Iowa Assessment and STAAR: