Math Recovery® Efficacy and Effectiveness Research

Pamela D. Tabor, Ph.D.[1]

The Institute of Education Sciences (IES) of the U.S. Department of Education and the National Science Foundation have developed a set of guidelines that articulate six types of educational research (*Common guidelines for education research and development*, 2013). These types are: 1) Foundational Research, 2) Early-Stage or Exploratory Research, 3) Design and Development Research, 4) Efficacy Research, 5) Effectiveness Research, and 6) Scale-up Research. The last three types contribute "to evidence of impact, generating reliable estimates of the ability of a fully-developed intervention or strategy to achieve its intended outcomes" (*Common guidelines*, 2013, p. 9). Efficacy studies are conducted on a limited scope under ideal conditions in which treatment delivery is conducted by highly trained individuals with typically limited diversity among the study participants. Effectiveness studies are conducted in less controlled, more authentic situations with a variety of populations delivered by a range of implementors. Efficacy studies ask the question, "Under ideal conditions, does this intervention perform as intended?" Effectiveness studies of educational interventions evaluate the extent to which an intervention successfully performs in real life contexts when delivered by typical educators, with a range of students.

Over the years, a number of studies have been conducted broadly evaluating the effectiveness of Math Recovery® intervention and Add+VantageMR® professional development for classroom teachers and its predecessor *Count Me In Too*. Although several evaluations have been conducted on the effect of Math Recovery® practice in a variety of implementations internationally (Graven, Stott, Mofu, & Ndongeni, 2015; Holliday, 2007, 2008; Willey, 2009; Willey, Holliday, & Martland, 2007), the purpose of this white paper is to provide an overview of research specific to the implementation within the United States of America in a chronological manner. The studies range from district-based internal evaluations and master's theses to an IES-funded randomized control trial (RCT) external evaluation across multiple states.

Program Development

In the mid-1990s Math Recovery® intervention was developed by Dr. Robert J. "Bob" Wright and his colleagues in New South Wales, Australia. As a part of the design research, diagnostic assessments and teaching procedures were trialed in the context of teaching experiments in a strictly controlled research setting. These assessments and teaching experiences were all videotaped, and the video was analyzed to determine the efficacy of various assessment protocols and teaching procedures. The highly efficacious assessments and procedures were codified into the Math Recovery® intervention assessment schedules and teaching procedures and were reported in *Early Numeracy: Assessment for Teaching and Intervention* (Wright, Martland, & Stafford, 2000, 2006) and *Teaching Number: Advancing Children's Skills and Strategies* (Wright, Martland, Stafford, & Stanger, 2002, 2006) now both in their second edition. These

[1] The author wishes to acknowledge the work of many in the Math Recovery® community who contributed to this body of research and reviewed drafts of this article. Bob Wright was particularly insightful in his comments and suggestions.

assessments and teaching procedures were designed for a highly intensive, one-on-one, non-scripted intervention detailed at length in the two volumes.

Almost immediately after the intervention was developed, adaptations for classroom teachers were developed in conjunction with the New South Wales Department of Education and Training (DET). This classroom teacher professional development was initially piloted as *Count Me In* and then scaled up as *Count Me In Too (CMIT)* across the state of New South Wales. Over the next decade and a half, the DET commissioned a number of external evaluations focusing on many different aspects of the CMIT program implementation (Bobis, 1996, 1997, 1999, 2001a, 2001b, 2003, 2004, 2006, 2009; Bobis et al., 2005; Bobis & Whitton, 1999). Overall, Bobis et al. found that CMIT was highly effective in changing teacher perceptions, beliefs, practices, and knowledge and improving student numeracy skills as measured on standardized state assessments. CMIT was later implemented and evaluated in New Zealand (Thomas & Ward, 2001) and the United States of America (MacLean, 2003).

Early Evaluation

Phillips, Leonard, Horton, Wright, and Stafford (Phillips, 1999; Phillips, Leonard, Horton, Wright, & Stafford, 2003) report on the evaluation of a pilot implementation of Math Recovery® intervention among first grade students in a public school district in upstate South Carolina. This quasi-experimental study used convenience samples to compare the growth of twelve students receiving the intervention with ten similar counterparts. The study analyzed data from the Math Recovery® assessments administered pre, post, and delayed post to participants and counterparts. The study results were characterized in this manner:

> At the beginning of the program, the control group was significantly higher in arithmetical strategies than was the Math Recovery group (*t*=-2.162, *p*=.043), but at the end of the intervention, the Math Recovery children significantly out-performed the control group (t=2.827, p=.01). Between the two groups, the two means are *t* standard errors different. The *p*-value is the probability that we would see such a value if there is no difference…Viewed another way, the experimental group made significant gains in arithmetical strategies (*t*=5.451, *p*=.000) during the period of the intervention and the control group did not (*t*=2.000, p=.343) (Phillips et al., p. 109).

MacLean (2003), in a pretest, posttest control group with matched sets quasi-experimental design conducted in one public school district, compared under-performing first grade students from three different groups. The sample included 38 sets matched on baseline performance. The first treatment group of low-performing students received Math Recovery® in schools in which the classroom teachers also participated in CMIT. In the case of the second group, the students did not receive the intervention, but their teachers participated in CMIT. The third group neither received the intervention nor did their teachers participate in the CMIT professional development; their math instruction was business as usual. MacLean compared the pre- to post-test demonstrated growth on six different aspects of the Learning Framework in Number. The analysis revealed no difference between the groups at baseline. Students in the Math Recovery® plus CMIT treatment group out performed students from the CMIT-only treatment as well as the control group (p<.001) with respect to stage growth, forward number word sequence, spatial patterning, and ten and ones (early base ten). The CMIT-only group was not statistically higher than the control group in these aspects. With respect to numeral identification, both the Math Recovery®

intervention plus CMIT and the CMIT-only treatment groups significantly (p<.001, and p=.017 respectively) outperformed the control group. MacLean concluded that the Math Recovery® intervention in conjunction with the professional development of classroom teachers via CMIT led to statistically significant greater gains in early numeracy than the professional development alone or business as usual math instruction. One question that was not asked was, "How does the student intervention alone perform, without the accompanying professional development of and coordination with classroom teachers?"

NAIAHEDC Evaluation

Work began in the early 2000s to further develop the classroom adaptations of Math Recovery® into the American context. This was first branded as *Strength in Number* (SN) among indigenous populations under the auspices of the *National American Indian, Alaskan & Hawaiian Educational Development Center* (NAIAHEDC) [later renamed *The First People's Center for Education*] headquartered in Sheridan, Wyoming. SN was piloted in a school situated on a reservation with a 99% American Indian population. In 2002-03, prior to implementation, the school had 0% proficient in math in Grade 3 or 6 on the state assessment, Proficiency Assessments for Wyoming Students (PAWS). During the 2004-05 school year, awareness sessions about Math Recovery® were conducted. Math Recovery® Intervention Specialist (MRIS) professional learning was initiated during the summer of 2005 with math intervention teachers and SN began a rolling implementation in the fall of the 2005-06 school year with Pre-K, K and Grade 1 staff. Grade 2 staff began SN in February of 2006. In 2006-07, two years into commencing Math Recovery® and SN professional learning with the staff, 78% of third graders and 40.54% of sixth graders tested proficient on PAWS (Education, 2007; WIS, 2009).

During 2007-08 and 2008-09 the NAIAHEDC implemented their Math Recovery® and SN Comprehensive Mathematics Program in additional schools serving Native populations. The program involved a five-tiered model of implementation. Tier 0 involved Math Recovery® awareness sessions and communication of data. Tier 1 involved the establishment of a Math Recovery® Interventionist and a steering committee with textbook review. Tier 2 added SN implementation and textbook adoption. Tier 3 continued with SN implementation and added in continuing support via collaborative teams. Tier 4 moved the implementation into a site-led, self-sustaining model. Based on initial internal evaluation positive results, NAIAHEDC expanded efforts into three additional states.

The evaluation investigated whether there was an increase in student achievement in mathematics among Native students in schools with significant Native populations. This was investigated on both a formative and summative level. With respect to formative goals, the external evaluator found:

> …70.5% (31 of 44) of 2009-2010 Math Recovery students made a growth of at least two stages. [Source: NAIAHEDC MR Data.] The data were coded as a 1 if the target was met and a 0 if the target of at least two stages of growth was not met. This distribution was tested against a binomial distribution (based on a Z approximation) with a test probability of .05. The mean was .70 and the standard deviation was .462. This distribution is at a highly significant level (p=.01). The chances of the scores being attributable to chance are extremely remote (Tabor, 2010, p. 20).

Furthermore, on more distal, summative assessments:

> …when analyzing the growth from the lowest quartile to higher quartiles, it is reasonable to consider any student in a proficiency level of 2 or higher as having demonstrated growth from the lowest quartile [at baseline] to a higher quartile. 14 out of 16 (87.5%) students demonstrated such growth [of at least one quartile shift higher]. This is a highly significant finding (Mean=.88 and Standard Deviation=.342, *p*=.004) (Tabor, 2010, p. 20).

The evaluator concluded that Math Recovery® and SN contributed to the improved success of the students over a three-year period as a part of a comprehensive school improvement plan on both proximal and more distal, standardized assessments of mathematics.

A 2011 follow-up study at one of the original implementation schools found that former Math Recovery® intervention participants out-performed all students in the school on the 2011 PAWS assessment (MacCarty, 2019).

In 2009, NAIAHEDC partnered with two borough school districts in Alaska in the KINSMEN Project to strengthen mathematics instruction and learning among largely Native student populations spanning diverse geographical areas. Classroom teachers participated in SN while specialists undertook Math Recovery® Interventionist professional learning. All teachers took a "12-item, multiple-choice assessment" (*KINSMEN Project final evaluation report*, 2013, p. 5) focusing on teacher knowledge of the "Learning Framework in Number and its application to teaching and learning" (*KINSMEN Project final evaluation report*, 2013, p. 5). A sampling of 26 participants' pre and post data was collected. "Pre and post comparison of the data indicate growth in content knowledge for SN's Learning Framework in Number" (*KINSMEN Project final evaluation report*, 2013, p. 5). The pre-to-posttest mean grew from 26% to 79% with an average gain of 6.4 items out of 12 (*KINSMEN Project final evaluation report*, 2013).

While SN was being implemented in schools with significant Native populations, a similar program of professional development was branded under the auspices of the US Math Recovery Council® (USMRC) as Add+VantageMR® designed for a more general cross section of the U.S. population. Add+VantageMR® was first published circa 2005.

IES Evaluation

During the mid-2000s, the Institute of Educational Sciences (IES) of the U.S. Department of Education funded an opportunity to investigate the effectiveness of Math Recovery® intervention without the typical Add+VantageMR® professional learning provided to the classroom teachers. This was a Randomized Control Trial (RCT) that strictly controlled for any cross contamination. In order to make sure none of the control students were exposed to any Math Recovery® instructional practices, the classroom teachers did not participate in the Add+VantageMR® professional development and the Math Recovery® interventionist was not allowed to communicate with the classroom teacher about what the student was learning in the intervention in any way. This is obviously not the standard operating procedure when Math Recovery® is introduced in a school; however, the researchers deemed it necessary in order to have a clean, "gold standard" randomized trial. The USMRC would never recommend this form of implementation. In Math Recovery® business as usual, the classroom teachers undertake Add+VantageMR® professional development at the same time as the specialists do Math Recovery® intervention; whereas, in the IES study, the classroom teachers did not do Add+VantageMR® when the

specialists were undertaking Math Recovery®. Nevertheless, it is sometimes necessary to isolate components to scrutinize their effectiveness.

Smith, Cobb, Farran, Cordray, Munter and Dunn reported some fairly impressive findings during the first year of implementation:

> The first year results show a small to moderate effect of participation in MR on [Woodcock-Johnson III (WJ III)] scores and moderate to large effects on the MR proximal assessments. Specifically, differences in the end of first grade mean scores on the WJ III subtests between students selected for tutoring and those on the waitlist ranged in effect size from .21 on the quantitative concepts scale to .28 on the applied problems scale (all differences statistically significant at the $p<.05$ level). Effect sizes on the MR 1.1 screening assessment ranged from .34 on the forward number sequence scale to .92 on the arithmetic strategies measure. These results compare favorably to those reviewed recently by Slavin and Lake (2006), including several cooperative learning programs that had median effect sizes of at least +0.30 in studies using randomized experimental or randomized quasiexperimental designs, including Class wide Peer Tutoring (.33), Student Team Learning (.19-.60), and TAI Math (.28-.38) (Smith et al., 2010, pp. 4-5).

These data include students outside the target range of the intervention, that is, above the first quartile at pretest. When analysis was restricted to individuals below the 25th percentile on the pretest (the Math Recovery® intervention's identified target population), the distal effect sizes ranged from 0.30 – 0.40. When adjustments were made for pretest measures, the researchers found "… a predicted effect size between .4 and .5 for those in the lowest quartile of study participants at pretest..." (Smith et al., 2013, p. 417).

In a resource published by IES to facilitate the interpretation of effects of educational interventions, Lipsey and others (2012) established guidelines for interpreting educational effect sizes based on a survey of educational research literature. The authors analyzed effect sizes for RCTs and quasi-experimental studies for elementary, middle, and high school settings on both distal, standardized assessments as well as researcher developed proximal measures. They found a median effect size of elementary level RCTs on distal measures (Broad Scope) (n=230) as .07 and a mean of .08 (SD=.27). For standardized tests with a narrow scope such as selecting subsets of the entire battery (n=374), they found a median of .17 and a mean of .25 (SD=.42). For researcher-developed proximal measures, they found a median of .34 and a mean of .40 (SD=.55). Smith et al. (2013) estimated the effect size of .4 to .5 on the narrow, standardized measure (see pp. 416-417). (See also Smith, Cobb, Farran, Cordray, Munter, et al., 2010.) The upper end of this estimate is twice the mean scores found for that class of research in the literature. When looking at the proximal measure, Smith et al. (2013) report an effect size for the entire sample of 1.04 which is 2.6 times the mean of .40 and well exceeding a standard deviation above the mean for the same class of measure from the Lipsey et al. (2012) IES report.

Smith et al. (2013) also concluded that students with lower pretest scores were more likely to benefit from Math Recovery®. However, researchers failed to find that the rate of gains was maintained at the end of second grade among the students who received the intervention during Year 1 (see Smith et al., 2013, p. 422). Smith et al. (2013) did not report any longitudinal analysis for the restricted subset that represents the target population (i.e. students in the first quartile at pretest); therefore, speculation would be necessary to determine what the odds of finding significance among the Math Recovery® target

subset would be, given the dramatic difference at the end of grade 1 between the target population and the entire data set. No follow-up data were gathered nor analyzed for Year 2 students at the end of second grade. Since *n* size greatly impacts significance testing (Maxwell & Delaney, 2004, p. 98ff) and the analysis of the target population was not disaggregated in the Year-2 follow-up data, it is not surprising that the study failed to find a sustained growth trajectory at a statistically significant level. It should be noted that Smith et al. recognized the weakness of their study that did not attend to the nature of the classroom instruction in either first or second grade classes.

One question that needs to be investigated is whether there are certain pedagogical methods within the classroom that provide better support for the successfully discontinued Math Recovery® intervention students. Do some styles of classroom instruction "undo" the intervention gains? This question has not yet been formally investigated with respect to Math Recovery® intervention. However Clements, Sarama, Wolf and Spitler (2013), in a clustered randomized trial comparing the impact of subsequent instruction on maintaining gained effects of early intervention on longitudinal student assessment scores, found "…that the follow-through component is important for maintaining the learning trajectory engendered by the pre-K intervention. Without the follow-through component, the effects are smaller each year" (Clements et al., 2013, p. 26). In a follow-up study Watts, Duncan, Clements and Sarama (2017) conclude:

> Our pattern of results has implications for developmental theory. If our fifth-grade finding is found to be robust to replication, then this would suggest that skill-building processes do not necessarily unfold in a monotonic manner. In other words, early math skills might not reliably lead to the development of later mathematical knowledge across all settings. Rather, early mathematical knowledge may only lead to the production of later knowledge when this early knowledge base is paired with the correct mix of content and teaching. This suggests that subsequent environments play a critical role in sustaining cognitive development in the wake of early investments in cognitive skills. This also suggests that skill-building theories that predict that early knowledge gains will necessarily lead to advantages in later achievement (e.g., Cunha & Heckman, 2008) may need some revision, as our results imply that skill development may be a more complex process that relies on many factors other than the mere possession of early skill advantages (Watts et al., p. 14).

The results of the Math Recovery® IES-funded evaluation study and the research of Clements and his colleagues suggest that intervention in isolation divorced from the classroom practices is not sufficient to maintain the higher growth trajectory longitudinally after the intervention is completed. When one considers the results of MacLean (2003) and Smith et al. (Munter, 2014; Munter, Garrison, Cobb, & Cordray, 2010; Smith, Cobb, Farran, Cordray, Munter, et al., 2010; Smith, Cobb, Farran, Cordray, & Munter, 2010, 2013), it seems reasonable to conjecture that for low-attaining students to accelerate their learning and maintain those new growth trajectories, a combination of low-attaining students receiving Math Recovery® intervention while their classroom teachers are implementing strategies learned in Add+VantageMR® has the highest likelihood of long term success.

Kentucky Center for Mathematics Evaluation

While the IES grant was being conducted with rigorous "gold standard" methodology that is not practical or even ethical for most evaluation efforts, an external evaluation was conducted on behalf of the Kentucky Center for Mathematics (KCM) by the University of Cincinnati Evaluation Services Center. The KCM is a state-funded initiative established in 2006 to improve mathematics instruction and learning across the state. The evaluation used convenience samples of the student Terra Nova CAT6 Version 2 results of schools that elected to implement one of two different math interventions, one of which was Math Recovery® intervention. CAT6 was administered, pre and post to all intervention students with a subset of 59 matched pairs from each intervention. Average grade level gains were calculated for the participants of each intervention. While the average growth for both interventions exceeded a year's growth, the analysis of matched pairs revealed that the MR students made an average gain of 2.22 grade levels growth in one school year while students from the other intervention made an average of 1.56 years growth. Furthermore, when looking at the percentage of students who reached grade level expectations by the end of the year for the entire sample, 70 of the MR students achieved grade level expectations whereas only 29% of the students from the other intervention achieved grade level expectations even though they had initially scored higher at pre-test than the MR participants (Ludwig, Jordan, Maltbie, & Marks, 2007). The gains of the Math Recovery® intervention students so out-performed the other intervention students that KCM decided to augment the professional learning of the other intervention program's teachers with Add+VantageMR® professional development in an effort to improve student growth among the schools implementing the other intervention (Gabbard, 2008). The next year's evaluation did not find nearly as dramatic a difference between the student growth of the two interventions, thus validating their additional Add+VantageMR® professional learning of the teachers from the other intervention program (Ludwig, Jordan, & Maltbie, 2008). It is significant to note that after a decade of implementation, Math Recovery® and Add+VantageMR® are still central to their professional learning in their Primary Mathematics Program (KCM, 2012), in the Kentucky Numeracy Project (KNP) website (KCM, 2014), the K Plus (Kentucky Numeracy Project Intensive Plus) project (KCM, 2015), and the Kentucky Numeracy Project Intensive Course (KCM, 2018). During the 2017-18 school year, 92 educators participated in the Kentucky Numeracy Project Intensive Course which included Math Recovery® professional learning as a part of the project (KCM, 2018).

Michigan Grant Evaluation

As part of a five-year, state-wide, grant-funded initiative in Michigan in which 228 classroom teachers in Phase I and 279 in Phase II participated in Add+VantageMR® professional learning course 1 and course 2, 19 district-level teacher-leaders became Math Recovery® Intervention Specialists to support classroom teachers participating in Phase I and Phase II Add+VantageMR® professional learning, and an additional 14 regional leaders underwent Math Recovery® Intervention Specialist professional learning in order to become regional trainers in new regions of Michigan. Extensive external program evaluations were conducted by the Science and Mathematics Program Improvement (SAMPI) at Western Michigan University. The SAMPI produced four comprehensive annual reports and numerous evaluation question-specific reports (Ruhf, Everett, & Miller, 2017; Ruhf et al., 2016a; Ruhf et al., 2016b; Ruhf, Miller, & Witucki, 2017; Ruhf & Sydlik, 2016a, 2016b; Ruhf et al., 2015; Stein, Everett, Ruhf, & Sydlik, 2016; Williams, Ruhf, Sydlik, & Lew, 2016). These evaluations were multi-faceted and involved analysis of both teacher

and student level change. Teacher level data included both survey and observational data as well as mathematics content pre- to posttest teacher data. Student level data included pre-post mathematics achievement data on a SAMPI-designed, validated mathematics content instrument as well as Rasch unit RIT data from the NWEA Mathematics Measure of Academic Progress (MAP).

With respect to the impact of Add+VantageMR® on teachers, SAMPI made several findings based on teacher surveys, teacher interviews, a teacher mathematics content knowledge assessment, and a video assessment:

1) "Ninety-eight percent (98%) of the Cohort 1 teachers felt the program met their expectations" (Ruhf et al., 2015, p. 2). "Nearly all [Add+VantageMR®] teachers (97%) felt their knowledge of early numeracy was enhanced by the program. Many teachers learned new ways and tools to teach math and better understand how children learn and how to help them move forward in their learning" (Stein et al. 2016, p. 2).
2) All teachers (100%) said their ability to teach mathematics content changed or was enhanced by the program. Teachers stated that they now have more resources, games and manipulatives to support student learning and have a better understanding of how to teach math" (Stein et al., 2016, p. 3).
3) Add+VantageMR® "[t]eachers' perceptions of their knowledge of how children learn math (p=<.001) [sic.], their ability to assess student learning of math (p = <.001) [sic.], and their ability to differentiate instruction for students (p= <.001) [sic.] significantly improved across both cohorts. Their familiarity with Common Core standards [sic.] also improved significantly" (Williams et al., 2016, p. 3).
4) With the first cohort, SAMPI found Add+VantageMR® participation is associated with significantly (p-value = 0.001) increased teacher (n=87) content knowledge in mathematics (Ruhf et al., 2015). Across both cohorts, SAMPI found a statistically significant ($p \leq .05$) association between teachers' (n=198) participation in the Add+VantageMR® professional development and pre-to-post growth in teacher content knowledge as measured by an instrument designed and validated by the external examiners (Ruhf & Sydlik, 2016a).
5) SAMPI further found statistically significant (p<.001) pre-to-post positive change in grades K-5 Add+VantageMR® participating teachers' ability in suggesting appropriate instructional strategies for each video vignette case presented (Ruhf & Sydlik, 2016b).

Findings with respect to student-level growth were equally impressive. In an analysis of a SAMPI-developed and validated assessment, students (n=2,869) from Add+VantageMR® classrooms (n=158) demonstrated statistically significant pre-to-post growth (α<0.05) in every grade level (Ruhf et al., 2016). They further found that among 438 students of 22 Add+VantageMR® teachers, students demonstrated statistically significant (α=0.05) pre-to-post growth on NWEA Mathematics Measure of Academic Progress (MAP) RIT (Rasch Unit) scores "…and exceeded the projected growth for all grade levels" (Ruhf et al., 2017, p. 2). Furthermore, both struggling and high-achieving students made significantly more growth on the nationally normed test when their teachers had participated in Add+VantageMR® professional learning (Ruhf et al., 2017).

Massachusetts Evaluation

Based on the very encouraging results from other regions, in the fall of 2017 the state of Massachusetts began the state-wide, grant-funded Massachusetts Tiered System of Support (MTSS) Tiered Math Academy project using Math Recovery® Intervention Specialist and Add+VantageMR® professional learning courses to build strong mathematics instructional leaders and implement a MTSS in struggling schools and districts. The Academy offers Add+VantageMR® Course 1, Add+VantageMR® Course 2, Add+VantageMR® Fractions, and Math Recovery® Intervention Specialist as a suite of professional learning opportunities implemented over a three-year cycle. Based on its internal evaluation, the Department of Education extended the contract with the USMRC, forming a second cohort increasing the number of districts and schools participating (MDEOASE, 2019). As of spring semester of 2019, the grant has impacted fourteen different school districts, in 77 schools. A total of 336 teachers have participated in the different professional learning opportunities: 215 teachers have completed Add+VantageMR® Course 1 and 2, 73 teachers have completed the Add+VantageMR® Fractions, and 48 teachers have completed Math Recovery® Intervention Specialist professional learning. Calculating 336 teachers times an average of 25 students per teacher estimates 8,400 students impacted by the grant in Year 2 alone. In actuality, 8,400 is likely a conservative annual student impact figure because six middle and high schools were among the participating schools. Secondary teachers typically teach more than 25 students a year. Many of the districts have elected to develop an in-district Add+VantageMR® Champion, enabling the district to offer the Add+VantageMR® courses within district to additional teachers so that their ongoing professional learning efforts will be sustainable beyond the scope of the MTSS grant.

Facilitating Local Evaluations

Since before the founding of USMRC, many local districts have conducted internal evaluations in the attempt to determine the effectiveness of their Math Recovery® and Add+VantageMR® implementations. While it is not possible to share the findings of every district-level evaluation in the scope of this paper, a few such efforts will be highlighted to demonstrate the types of analyses that have been conducted.

Richmond County Schools

Richmond County Schools local educational authority in North Carolina was one of the early implementors of Math Recovery® intervention in the United States. In a follow-up study of former MR participants, Gibson (2001) found a correlation between students' exit arithmetical stage as first grade intervention students and their future performance as third and fourth grade students. Students who had attained a post-assessment stage of 3 or higher (the MR target for successful completion), met end of grade expectations as third and fourth grade students.

Roye-Williams Elementary School

Mathematics Recovery® was implemented at Roye-Williams Elementary School (RWES), a Title 1 school in Harford County, Maryland during the 1999-2000 school year. One Mathematics Recovery® teacher and three para professionals were dedicated full-time to mathematics instruction and professional development. Prior to the implementation of Mathematics Recovery®, 30-40% of children scored

proficient or better in mathematics on the Maryland state assessment program (MSPAP). This was well below the district average. In 2004, the fifth year of implementation, 76.6 % of fifth-grade students were proficient or better in mathematics on the Maryland State Assessment (MSA). 82.9% of fourth-grade students scored proficient or better with 22.9 % scoring advanced. 72.8 % of the third-grade students scored proficient or better. If one analyzes the disaggregated data, it is obvious that the African American students, Hispanic students, and students receiving special services at RWES significantly outperformed their peers in both the district and state levels for that year's state assessment. This provides evidence that MR coupled with imbedded professional development centered around the MR Learning Framework in Number is effective in closing the achievement gap among traditionally underperforming groups.

Table 1. RWES Longitudinal Data: Grade 5

No Child Left Behind Disaggregated Data for 2004 5th Graders			
Category	% of Children Proficient or Better		
	RWES	Harford Co.	Maryland
Africa American	75.0%	54.8 %	45.9%
White	75.7%	78.5%	76.3%
Hispanic	85.7%	70.3%	52.1%
Limited English Proficient	76.3%	65.0%	36.0%
504 Services	76.6%	81.0%	59.1%
Free & Reduced Lunch	60.7%	56.1%	44.4%
Special Education	41.7 %	40.8%	29.7%

http://www.mdk12.org/data/msa_data/index.asp?K=120639

Table 2. RWES Longitudinal Data: Grade 4

No Child Left Behind Disaggregated Data for 2004 4th Graders			
Category	% of Children Proficient or Better		
	RWES	Harford Co.	Maryland
Africa American	77.5 %	57.9 %	51.9 %
White	83.3 %	84.0 %	83.2 %
Hispanic	91.7 %	73.7 %	59.3 %
Limited English Proficient	NA	45.0 %	38.8 %
504 Services	NA	68.7 %	63.9 %
Free & Reduced Lunch	76.9 %	61.1 %	51.4 %
Special Education	62.5 %	48.5 %	38.5 %

http://www.mdk12.org/data/msa_data/index.asp?K=120639

Table 3. RWES Longitudinal Data: Grade 3

No Child Left Behind Disaggregated Data for 2004 3rd Graders			
Category	% of Children Proficient or Better		
	RWES	Harford Co.	Maryland
Africa American	75.9 %	67.1 %	51.3 %
White	76.3 %	84.0 %	64.5 %
Hispanic	50.0 %	68.7 %	58.2 %
Limited English Proficient	NA	50.0 %	49.9 %
504 Services	NA	100 %	62.6 %
Free & Reduced Lunch	75.0 %	65.3 %	55.9 %
Special Education	44.4 %	51.1 %	42.1 %

http://www.mdk12.org/data/msa_data/index.asp?K=120639

Green Bay Area Public School District

Prior to implementation in 2014, Green Bay, Wisconsin was a "clean" site, having no previous exposure to Math Recovery® or Add+VantageMR®. In the spring of 2014 the district started by training 100% of their staff in two Title I buildings. Due to the staff's overwhelmingly positive response to their learning, the district invested in training 17 MRIS and approximately 350 classroom teachers in Add+VantageMR® 1 and 2 in the summer of 2014. After completion of MRIS certification, the 17 MRIS undertook Add+VantageMR® Champion professional learning so the district would have a Champion in each of their neediest schools.

In order to evaluate the impact of this professional learning, Mathematics Teaching Efficacy Beliefs Instrument (MTEBI) (Enochs, Smith, & Huinker, 2000; Riggs, Fischman, Riggs, Jetter, & Jesunathadas, 2018) was administered pre and post to all participating teachers. The evaluation found teachers' participation in Add+VantageMR® professional development positively impacted teachers' beliefs that effective teaching of mathematics can bring about student learning regardless of external factors (outcome expectancy) (Miller, 2019). The data from this initial group were from "clean" teachers and involve repeated measures (pretest gathered prior to participation in Add+VantageMR® course 1 and posttest gathered after completion of Add+VantageMR® Course 2). The sample in the analysis was restricted to elementary level teachers (PreK-Grade 5) who had completed both courses.

In a dependent sample t-test (n=129 elementary teachers pre to post), there was a statistically significant positive change in post test scores for the Outcome Expectancy Subset over pretest (p = .000). See Tables 4 and 5 below for details.

Table 4. Green Bay Data 1

One-Sample Statistics Outcome Expectancy Subset

	N	Mean	Std. Deviation	Std. Error Mean
Pre 1 Score - OE	129	28.25	3.493	.308
Post 2 Score - OE	129	29.38	3.700	.326

Table 5. Green Bay Data 2

One-Sample Test Outcome Expectancy Subset

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Pre 1 Score - OE	91.842	128	.000	28.248	27.64	28.86
Post 2 Score - OE	90.185	128	.000	29.380	28.74	30.02

Furthermore, the Self Efficacy subset, designed to measure a teacher's belief in his or her own ability to teach mathematics (Enochs et al., 2000) also demonstrated a highly significant positive change from pretest to post (*p*=.000). See Tables 6 and 7 for the details.

Table 6. Green Bay Data 3

One-Sample Statistics SE Subset

	N	Mean	Std. Deviation	Std. Error Mean
Pre 1 Total Score - SE	129	51.63	5.353	.471
Post 2 Score - SE	129	51.88	4.614	.406

Table 7. Green Bay Data 4

One-Sample Self Efficacy Test Subset

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Pre 1 Total Score - SE	109.537	128	.000	51.628	50.70	52.56
Post 2 Score - SE	127.722	128	.000	51.884	51.08	52.69

The total composite score for both subsets also demonstrated a statistically significant difference from pretest to post (*p*=.000). See Tables 8 and 9 for the details.

Table 8. Green Bay Data 5

One-Sample Statistics - Total Score

	N	Mean	Std. Deviation	Std. Error Mean
Pre 1 Total Belief Test	129	79.89	6.734	.593
Post 2 Total Belief Test	129	81.26	6.335	.558

Table 9. Green Bay Data 6

One-Sample Test - Total Score

Test Value = 0

	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Pre 1 Total Belief Test	134.747	128	.000	79.891	78.72	81.06
Post 2 Total Belief Test	145.700	128	.000	81.264	80.16	82.37

Based on these positive findings, it is not surprising that the Green Bay Area School District's training of new teachers in Add+VantageMR® and investing in MRIS and Add+VantageMR® Champions is currently ongoing, even through the transition of key district mathematics leadership.

Solon City School District

The Solon City School District in Ohio conducted a five-year longitudinal follow-up of students with MR intervention students.

> Twenty of twenty-six fifth grade students who received Math Recovery® services in their first-grade year, had data from their state exams in third and fourth grades. Of those 20 students, 95% were proficient, accelerated, or advanced as third graders. 75% were proficient, accelerated, or advanced as fourth graders. Twenty-nine of thirty-four fourth grade students who received Math Recovery® services in their first-grade year, had data from their state exams in third grade. 86% received proficient, accelerated, or advanced scores on their third-grade [state] exams (Silvestri, 2015).

Supporting Future Local Evaluation Efforts

The USMRC so believes that program evaluation is important that the Board of Directors has designed research and evaluation as a strategic initiative for the past two years. As a part of this effort, a session was provided at the USMRC Conference entitled, "Nuts and bolts of evaluating your Math Recovery® Add+VantageMR® implementation: A constructive work session" in which participants received an overview of evaluation strategies and then were given time and planning tools to begin developing a plan for their internal evaluations. The aim of the session was to empower school, district and regional math leaders with the knowledge to more effectively evaluate their own implementations. The USMRC plans to make similar presentations at state and regional conferences. While the USMRC is not privy to all local evaluations, the Math Recovery® community has been encouraged to share their findings with the USMRC. Spokane Public School District has begun a local, internal evaluation utilizing the information gained from these USMRC presentations. Initial analysis is very promising, and they hope to have results

to share publicly in the near future (Dodd, 2019). The USMRC is also in the planning stages for developing MR Connect Evaluation Corner for the members only section of the website. This section will include planning tools, resources, exemplars, frequently asked questions, and a discussion board to support local and state administrators in their evaluation efforts.

Ongoing USMRC Evaluations Efforts

As a part of the ongoing evaluation research efforts of the USMRC, it has contracted with the nationally-recognized, independent research firm, HumRRO, to conduct an external evaluation of Add+VantageMR Course 1 and Course 2. This three-year study began in the fall of 2018 and will involve a comparative analysis of student standardized test scores from Add+VantageMR® implementation schools compared to matched non-implementation schools within one district in the western United States. The study will also be analyzing the nature of the change in teacher knowledge and practices and school culture that accompany Add+VantageMR® professional learning. Strategic plans are also being drawn for a future, larger-scale external evaluation.

Efficacy and effectiveness research are an ongoing priority of the USMRC. Considerable resources, both human and capital, are being dedicated to this continuing effort. This white paper is a dynamic document that will be amended to reflect the USMRC's ongoing evaluation efforts.

References

Bobis, J. (1996). *Count Me In Too: 1996 Report: Report of the evaluation of the Count Me In Project*. Macarthur, AU: University of Western Sydney.

Bobis, J. (1997). *Count Me In Too: 1997 report*. Macarthur, AU: University of Western Sydney.

Bobis, J. (1999). *Count Me In Too: 1999 Report: The Impact of Count Me In Too on the professional knowledge of teachers*. Sydney, AU: New South Wales Department of Education and Training.

Bobis, J. (2001a). *Count Me In Too: 2000 Report: A case study of implementation: A report prepared on behalf of the NSW Department of Education & Training*. Sydney, AU: NSW Department of Educatin & Training.

Bobis, J. (2001b). *Count Me In Too: 2001 Report: The effect of Count Me In Too on Year 3 Basic Skills Test results: A report prepared on behalf of the New South Wales Department of Education and Training*. Sydney, AU: NSW Department of Education & Training.

Bobis, J. (2003). *Count Me In Too: Evaluation of Stage 2: A report prepared on behalf of the NSW Department of Education & Training*. Sydney, AU: NSW Department of Education & Training.

Bobis, J. (2004). *Coount Me In Too: An Evaluation of the facilitator model: A report prepared on behalf of the NSW Department of Education & Training*. Sydney, AU: NSW Department of Education & Training.

Bobis, J. (2006). *Count Me In Too: Factors affecting the sustainability and expansion of Count Me In Too: A report prepared on behalf of the NSW Department of Education & Training*. Sydney, AU: NSW Department of Educaiton & Training.

Bobis, J. (2009). *Count Me In Too: The Learning Framework in Number and its impact on teacher knowledge and pedagogy: A report on behalf of the NSW Department of Education & Training for the Count Me In Too program 2008*. Sydney, AU: NSW Department of Education & Training.

Bobis, J., Clarke, B., Clarke, D., Thomas, G., Wright, R. J., Young-Loveridge, J. M., & Gould, P. (2005). Supporting teachers in the development of young children's mathematical thinking: Three large scale cases. *Mathematics Education Research Journal, 16*(3), 27-57.

Bobis, J., & Whitton, D. (1999). *Count Me In Too: 1998 Report: The mathematical achievement and self-concept of kindergarten and year 1 children*. Retrieved from Macarthur, AU:

Clements, D. H., Sarama, J., Wolfe, C. B., & Spitler, M. E. (2013). Longitudinal evaluation of a scale-up model for teaching mathematics with trajectories and technologies: Persistence of effects in the third year. *American Educational Research Journal*. doi:DOI: 10.3102/0002831212469270

Common guidelines for education research and development. (2013). Retrieved from https://ies.ed.gov/pdf/CommonGuidelines.pdf

Dodd, B. (2019, January 25). [Personal email communication].

Education, U. S. D. o. (2007). *Executuve Summary Grant Performance Report (ED 524B) for Grant Award # U215K040153*. Washington, D.C.

Enochs, L. G., Smith, P. L., & Huinker, D. (2000). Mathematics teaching efficacy beliefs instrument. *School Science and Mathematics, 100*(4), 194-202.

Gabbard, A. (2008). *Facilitating teacher growth for student success: Kentucky's primary mathematics diagnostic intervention initiative*. Retrieved from Highland Heights, KY: https://pdfs.semanticscholar.org/633f/7b70fbbf8d84383a18291a89ea8e1f8a46a3.pdf

Gibson, J. H. (2001). *Richmond County Schools: Title I: Math Recovert/Early Math Program: 2000-01*. Retrieved from Richmond County, NC:

Graven, M., Stott, D., Mofu, Z., & Ndongeni, S. (2015). *Identifying stages of numeracy proficiency to enable remediation of fundational knowledge using the Learning Framework in Number*. Paper presented at the 23rd Annual Conference of South African Association for Research in Mathematics, Science and Technology Education, Maputo, Mozambique. https://www.ru.ac.za/media/rhodesuniversity/content/sanc/documents/SAARMSTE%202015%20LFIN%20Graven%20Stott%20et%20al.pdf

Holliday, A. (2007). *Mathematics Recovery Evaluation*. Retrieved from Cumbria, UK:

Holliday, A. (2008). *Mathematics Recovery Evaluation Report*. Retrieved from Cumbria, UK:

KCM. (2012). KCM: Kentucky Center for Mathematics: Annual Report: 2012. Retrieved from https://www.kentuckymathematics.org/docs/2012AnnualReport.pdf

KCM. (2014). KCM: Kentucky Center for Mathematics: Annual Report: 2014. Retrieved from https://www.kentuckymathematics.org/docs/AnnualReportKCM2013-14.pdf

KCM. (2015). KCM: Kentucky Center for Mathematics: Annual Report: 2015. Retrieved from https://www.kentuckymathematics.org/docs/AnnualReportKCM2015.pdf

KCM. (2018). KCM: Kentucky Center for Mathematics: Annual Report: 2018. Retrieved from https://www.kentuckymathematics.org/docs/KCMAnnualReport-2018.pdf

KINSMEN Project final evaluation report. (2013). Retrieved from Sheridan, Wyoming:

Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., . . . Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms. (NCSER 2013-3000).* Retrieved from Washington, DC: https://ies.ed.gov/ncser/pubs/20133000/pdf/20133000.pdf

Ludwig, K., Jordan, J., & Maltbie, C. (2008). *Kentucky Center for Mathematics: 2007-2008 intervention program evaluation: Summary of results*. Retrieved from Cincinnati, OH:

Ludwig, K., Jordan, J., Maltbie, C., & Marks, J. (2007). *Kentucky Center for Mathematics: 2006-2007 evaluation: Summary of results*. Retrieved from Cincinnati, OH:

MacCarty, L. P. (2019, 1/30/2019). [Personal email].

MacLean, H. E. (2003). *The effects of early intervention on the mathematical achievement of low-performing first grade students.* (Doctor of Education Doctoral dissertation), University of Houston, Houston, TX. ProQuest Dissertations and Theses, AAT 3081498.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

MDEOASE. (2019). Multi-Tiered Systems of Support (MTSS) Academies 2018-2019. Retrieved from http://www.doe.mass.edu/sfss/prof-develop/academies-18-19.html

Miller, C. (2019, February 8). [Green Bay, WIsconsin Teacher Change Data].

Munter, C. (2014). Developing visions of high-quality methematics instruction. *Journal for Research in Mathematics Education, 45*(5), 584-635.

Munter, C., Garrison, A., Cobb, P., & Cordray, D. (2010). *Evaluating Math Recovery: Measuring fidelity of implementation*. Paper presented at the 2010 SREE Conference: Research into Practice, Washington, D.C. https://www.sree.org/conferences/2010/program/

Phillips, V. J. (1999). *The effects of a math recovery program: A comparative study.* (Doctor of Philosophy Mixed methods), Clemson University, South Carolina.

Phillips, V. J., Leonard, W. H., Horton, R. M., Wright, R. J., & Stafford, A. K. (2003). Can Math Recovery save children before they fail? *Teaching Children Mathematics, 10*(2), 107-111.

Riggs, I. M., Fischman, D. D., Riggs, M. L., Jetter, M. E., & Jesunathadas, J. (2018). Measuring teachers' beliefs in relation to teaching mathematics with mathematical practices in mind. *School Science and Mathematics, 118*, 385-395. doi:10.1111/ssm.12303

Ruhf, R., Everett, K., & Miller, P. (2017). *Analysis of Normative Data from NWEA Mathematics Measure of Academic Progress (MAP) Grades K-5, 2015-16*. Retrieved from http://www.mathrecovery.org/pdfs/how-it-works/Analysis-of-MAP-Data.pdf

Ruhf, R., Everett, K., Sydlik, M. A., Miller, P., Souers, T., Williams, C., . . . Stein, C. (2016a). *Annual Evaluation Report: April 2015 – June 2016*. Retrieved from http://www.mathrecovery.org/pdfs/how-it-works/MI-Annual-Math-Recovery-Evaluation-Report-2016.pdf

Ruhf, R., Everett, K., Sydlik, M. A., Miller, P., Souers, T., Williams, C., . . . Stein, C. (2016b). *Annual Evaluation Report: Making an impact with Intel(R) Math Extension (MI^{2}M Extension): A Michigan Matheamtics and Science Parnership (MSP) Grant: Muskegon Area ISD Regional Mathematics and Science Center*. Retrieved from

Ruhf, R., Miller, P., & Witucki, A. (2017). *Annual evaluation report: Supporting implementation of Math Recovery(R) professional development (phase II): A Michigan Mathematics and Science Partnership (MSP) Grant: Muskegon Area ISD Regional Mathematics Science Center: July 2016-September 2017*. Retrieved from Michigan:

Ruhf, R., & Sydlik, M. A. (2016a). *Pre/Post Teacher Content Assessment*. Retrieved from http://www.mathrecovery.org/pdfs/how-it-works/MI-Teacher-Pre-Post-Test-Cohorts-1-and-2-Report.pdf

Ruhf, R., & Sydlik, M. A. (2016b). *Pre/post teacher video assessment*. Retrieved from http://www.mathrecovery.org/pdfs/how-it-works/MI-Teacher-Pre-Post-Video-Cohorts-1-and-2-Report.pdf

Ruhf, R., Sydlik, M. A., Everett, K., Jenness, M., Stein, C., & Miller, P. (2015). *Annual report for the period April 1, 2015 to October 31, 2015*. Retrieved from http://www.mathrecovery.org/pdfs/how-it-works/MI-Math-Recovery-2015-Annual-Report.pdf

Silvestri, T. (2015, June 29). [Personal email communication on longitudinal MRIS followup analysis].

Smith, T. M., Cobb, P., Farran, D., Cordray, D., Munter, C., & Dunn, A. (2010). *Evaluating Math Recovery: Assessing the Causal Impact of Math Recovery on Student Achievement*. Paper presented at the

2010 SREE Conference: Research into Practice, Washington, D.C. https://eric.ed.gov/?id=ED514486

Smith, T. M., Cobb, P., Farran, D. C., Cordray, D. S., & Munter, C. (2010). *Does the Mathematics Recovery Program Improve Students' Mathematics Learning?* Paper presented at the 2010 IES Research Conference, National Harbor, Maryland. https://ies.ed.gov/director/conferences/10ies_conference/presentations.asp

https://ies.ed.gov/director/conferences/10ies_conference/slides.asp?ppt=smitht

Smith, T. M., Cobb, P., Farran, D. C., Cordray, D. S., & Munter, C. (2013). Evaluating Math Recovery: Assessing the causal impact of a diagnostic tutoring program on student achievement. *American Educational Research Journal, 50*(2), 397-428.

Stein, C., Everett, K., Ruhf, R., & Sydlik, M. A. (2016). *Cohorts 1 & 2 teacher interview report*. Retrieved from http://www.mathrecovery.org/pdfs/how-it-works/MI-Math-Recovery-Teacher-Interview-Final-Report.pdf

Tabor, P. D. (2010). *NAIAHEDC annual evaluation*. Retrieved from Havre de Grace, Maryland:

Thomas, G., & Ward, J. (2001). *An evaluation of the Count Me In Too Pilot Project*. Retrieved from Wellington, NZ: Available from Learning Media Customer Services, Box 3293 Wellington, NZ (Item # 10211).

Watts, T. W., Duncan, G. J., Clements, D. H., & Sarama, J. (2017). What is the long-run impact of learning mathematics during preschool? *Child Development*. doi:10.1111/cdev.12713

Willey, R. (2009). *A study of the impact of Mathematics Recovery training on teaching staff's constructs about the teaching and learning of number.* (Doctor of Educational Psychology), Univeristy of Newcastle upon Tyne, Newcastle, UK. (2160717200002411)

Willey, R., Holliday, A., & Martland, J. (2007). Achieving new heights in Cumbria: Raising standards in early numeracy through Mathematics Recovery. *Educational and Child Psychology, 24*(2), 108-118.

Williams, C., Ruhf, R., Sydlik, M. A., & Lew, M. M. (2016). *Post-program survey of teacher participants: Cohorts 1 and 2* Retrieved from http://www.mathrecovery.org/pdfs/how-it-works/MI-Math-Recovery-Teacher-Post-Survey-Report.pdf

WIS. (2009). *Celebrate our success 2008-2009: Wyoming Indian Schools*. Retrieved from Ethete, Wyoming:

Wright, R. J., Martland, J., & Stafford, A. K. (2000). *Early numeracy: Assessment for teaching and intervention*. London: Paul Chapman Publishing.

Wright, R. J., Martland, J., & Stafford, A. K. (2006). *Early numeracy: Assessment for teaching and intervention* (2nd ed.). London: Paul Chapman Publishing.

Wright, R. J., Martland, J., Stafford, A. K., & Stanger, G. (2002). *Teaching Number: Advancing children's skills and strategies*. London: Paul Chapman Publishing.

Wright, R. J., Martland, J., Stafford, A. K., & Stanger, G. (2006). *Teaching Number: Advancing children's skills and strategies* (2nd ed.). London: Paul Chapman Publishing.