A Synthesis of Middle School Mathematics Intervention Research for Students with Learning Disabilities

Sun A Kim, Kathleen Hughes, Jennifer Porterfield, Diane P. Bryant, Brian R. Bryant
University of Texas at Austin
Minyi Shih
California State University-Los Angeles
**Background: Significant Factors Influencing Mathematics Education/Special Education**

- National Mathematics Advisory Panel (2008) recommendations for curricular focus
- National Council of Teacher’s of Mathematics (2000) emphasis on problem solving
- National Council of Teacher’s of Mathematics (2006) *Curriculum Focal Points*
- Emerging importance of algebra as the “gatekeeper” for post-secondary opportunities
- High stakes tests focus on problem solving as the vehicle for assessing mathematics performance
- Learning characteristics of students with learning disabilities and difficulties in mathematics
Background: Purpose

- The purpose of this study was two-fold. First, we reviewed and synthesized the research on the mathematics performance of middle school/junior high school students with learning disabilities.

- Second, we examined the studies to determine the degree to which studies included information that satisfied the quality indicators for group (Gersten et al., 2005) and single-subject (Horner et al., 2005) designed research.
**Background: Questions that Guided this Synthesis**

- What mathematics domains are the focus of the mathematics interventions?

- What are the features and effects of interventions on the mathematics performance of middle school/junior high students with learning disabilities?

- To what extent do mathematics intervention studies satisfy quality indicators for group and single-subject designed studies?
Method: Selection Procedures

Step 1: Electronic Database Search
- ERIC, Psych Info, Dissertation Abstracts, Google Scholar
- Key words: Intervention, instruction, strategies, mathematics learning disabilities, mathematics learning difficulties, low achievers, mathematics, arithmetic, mathematics intervention, mathematics instruction, learning disabilities, instructional strategies, middle school students, and secondary education

Step 2: Ancestral Search
- Syntheses on mathematics interventions (e.g., Maccini et al., 2007) were examined to identify articles, reference list checking

Step 3: Manual Search

Step 4: Article Review
- Identify articles that meet inclusion criteria (done independently by authors; 100% agreement on articles to include)
- Confirm selection criteria: participants, design, intervention, outcome, publication
Method: Selection Criteria

• Participants
  • Middle School/Junior High School
  • Grades: 6th-9th
  • Learning disabilities

• Intervention
  • Procedure (e.g., strategy, instructional routine, CRA) employed to teach mathematical skills/concepts

• Research Designs
  • Single-subject, Experimental, Quasi-experimental

• Outcome/DV
  • Mathematics skills/concepts measured as the dependent variable

• Publication
  • Years: 2000-2008
  • Peer-reviewed journals
  • Studies conducted in the U.S.
Method: Coding Procedures and Inter-rater Reliability

• Descriptive Review
  • Researcher-developed coding form included information on
    • Participants [gender, age, ethnicity, disability, SES, etc.]
    • Settings
    • Research designs [experimental, quasi-experimental, single-subject]
    • Dependent variables/Mathematics domains
    • Independent variables [procedure, routine, duration/length]
    • Maintenance & generalization
    • Results

• Quality Indicators Review
  • Coding form included quality indicator categories from
    • Gersten et al., 2005. Quality indicators for group experimental and
      quasi-experimental research in special education. Exceptional
      Children, 71(2), 149-164.
    • Horner et al. (2005). The use of single-subject research to identify
      evidence-based practice in special education. Exceptional Children,
      71(2), 165-179.
Method: Coding Procedures and Inter-rater Reliability

• Procedures
  • The authors independently coded a single article, inter-rater reliability computed as percentage of agreement (agreements divided by agreements plus disagreements), inter-rater reliability average was 85%, consensus was reached on disagreements before the next step
  
  • Articles were assigned among the authors where each article was double-coded by two raters, inter-rater reliability average was 89% for single-subject studies and 79% for group studies, consensus was reached on disagreements before information was reported
Method: Effect Size (ES) Calculation

• Single Subject Design Studies
  - Within each study, PND was calculated for each student and averaged for the mean percent of non-overlapping data points (PND) for the study.
  - Note 1. Effective- PND > 69; Questionable - PND = 50 - 69; Ineffective- PND <50 (Scruggs & Mastropieri, 2001)

• Group Design Studies
  - Within each study, ESs were calculated for each measure and were averaged to produce the mean effect size for the study.
  - Posttest (and/or maintenance) effect size was calculated by adjusting Cohen’s d for treatment-control posttest scores using pretest performance (Shadish & Haddoc, 1994, cited in Jitendra & Sood, 1997).
  - Note 2. Large effect - d = .80; Moderate effect - d = .50; Small effect - d = .2 (Cohen, 1988)

\[
\text{Adjusted ES} = \frac{M_{e2} - M_{c2}}{SD_{pooled2}} - \frac{M_{e1} - M_{c1}}{SD_{pooled1}}
\]
Results: Participants

Single Subject Design Studies (n = 5)

- Total: 16 (male, n = 13, female, n = 3)
- Ages: 13 (n = 7), 14 (n = 6), 15 (n = 2), 16 (n = 1)
- Grades: 7th (n = 1), 8th (n = 13), 9th (n = 1), and 10th (n = 1)
- Ethnicity: African American (n = 4), Caucasian (n = 11), Hispanic (n = 1)
- Socioeconomic Status: Below Average (other terms included low, free/reduced lunch) (n = 5), Medium (n = 2), High (n = 4), not reported (n = 5)
- Disability Categories: Learning Disability (n = 13), Learning Disability and Emotional Behavioral Disturbance (n = 2), and Learning Disability and Speech and Language Disability (n = 1)
Results: Participants

*Group Design Studies (n = 4)*

- Total: 268 (male, n = 98, female, n = 102, not reported, n = 68)
- Ages: Range of 11 to 15 (specific ages not reported in 2 studies)
- Grades: 6th (n = 51), 7th (n = 201), 8th (n = 16)
- Ethnicity: Reported for 2 studies and only for students with disabilities in one of those studies: African American (n = 4), Caucasian (n = 19), Hispanic (n = 12)
- Socioeconomic Status: 27 “at-risk” students in Witzel et al. (2003) were from low-socioeconomic backgrounds, not reported in other studies
- Disability Categories: LD (n = 106), LD/ADD (n = 2), LD/ADHD (n = 1), LD/EBD (n = 11), LD/SL (n = 2), ADD (n = 1), EBD (n = 1), Other disabilities (n = 8)
# Results: Question #1

<table>
<thead>
<tr>
<th>Math Domain</th>
<th>#</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fractions</td>
<td>3</td>
<td>Bottge et al. (2007), Butler et al. (2003), Joseph et al. (2001)</td>
</tr>
<tr>
<td>Decimals</td>
<td>1</td>
<td>Bottge et al. (2007)</td>
</tr>
<tr>
<td>Percent</td>
<td>1</td>
<td>Bottge et al. (2007)</td>
</tr>
<tr>
<td>Algebra</td>
<td>1</td>
<td>Bottge et al. (2007), Maccini et al. (2000), Witzel et al. (2003)</td>
</tr>
<tr>
<td>Data Anal. &amp; Prob.</td>
<td>1</td>
<td>Bottge et al. (2007)</td>
</tr>
</tbody>
</table>
## Results: Question #2 (Single-subject Design Studies)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Math Domain</th>
<th>Intervention/Instructional Routine</th>
<th>Duration/ Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cass, Cates, Smith, &amp; Jackson (2003)</td>
<td>word problem-solving/area &amp; perimeter</td>
<td>- manipulative instruction&lt;br&gt;- mastery = 80%</td>
<td>modeling, prompted guided practice, independent practice&lt;br&gt;7-8 days for perimeter; 6-7 days for area/daily</td>
</tr>
<tr>
<td>Jitendra, DePipi, &amp; Perron-Jones (2002)</td>
<td>word problem-solving/multiplication &amp; division</td>
<td>- schema-based: 2 conditions-problem schemata id. (PSID) problem solution (PS)&lt;br&gt;- mastery = PSID &amp; PS: 100% for 2 sessions each</td>
<td>On average, PSID: 6 sessions, PS: 12 sessions/35-40 minutes per session</td>
</tr>
<tr>
<td>Joseph &amp; Hunter (2001)</td>
<td>fractions/computation problems</td>
<td>- cue-card strategy, 3 problem-solving instances with numerical representations&lt;br&gt;- mastery = 90-100%</td>
<td>teacher demonstration, student practice with prompt&lt;br&gt;27 sessions</td>
</tr>
<tr>
<td>van Garderen (2007)</td>
<td>word problem-solving/whole # computation/1- &amp; 2-step problems</td>
<td>- instruction for generating and using schematic diagrams&lt;br&gt;- mastery = 75% 3 consecutive probes</td>
<td>modeling, guided practice, independent practice&lt;br&gt;35 minutes 2-4 times per week</td>
</tr>
</tbody>
</table>
## Results: Question #2 (Single-subject Design Studies)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Type of SS Design</th>
<th>Dependent Variables</th>
<th>Interv./ES (PND¹)</th>
<th>Main./ES (PND)</th>
<th>Gen.</th>
</tr>
</thead>
</table>
| Cass, Cates, Smith, & Jackson (2003) | multiple-baseline design across people and two behaviors (skills)                | - # of problems correct  
- type of errors  
- length of time to solve problem  
- duration of training session | •Area: 84.9%  
•Perimeter: 82.1% | •Area: 100%  
•Perimeter: 100% | •All students able to solve gen. problems using geo-boards |
- extent to which schema strategy used effectively | 82.5%                              | 100% (3 out of 4 students-initial probe) | •Pre-treatment: M=37%  
Post-treatment: M=100% |
| Joseph & Hunter (2001)           | multiple baseline                                                                | - # of fraction problems calculated correctly                                         95.9%                              | 100%                                | NA                                                              |
| Maccini & Ruhl (2000)            | multiple probe design across three participants                                   | - % of strategy use  
- % correct on problem representation (PR)  
- % correct on problem solution/answer (PS) | •Representati on: 66.6%  
•Solution: 94.3% | •Representati on: 66.6%  
•Solution: 100% | •Near gen:  
PR: M=73%  
PS: M=67%  
•Far gen:  
PR: M=29.3%  
PS: M=28.7% |
| van Garderen (2007)              | multiple probe across participants                                               | - diagram use (DU)  
- diagram form (DF) [pictorial or schematic] | 55.5%                              | 66.6%                             | •DU: M=100%  
•DF: M=88%  
•Test: M=33% |
### Results: Question #2 (Group Design Studies)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Math Content</th>
<th>Intervention/Instructional Routine</th>
<th>Duration/ Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottge, Rueda, Serlin, Hung, &amp; Kwon (2007)</td>
<td>&quot;Real world problems&quot;</td>
<td>- 2 Enhanced Anchored Instructions (EAI) Kim’s Komet Challenge (KKC) &amp; Fraction of the Cost Challenge (FCC)</td>
<td>Kim’s Komet for 13 days; Fraction of the Cost for 11 days</td>
</tr>
<tr>
<td>Butler, Miller, Crehan, Babbit, &amp; Pierce (2003)</td>
<td>word problems-solving/equivalent fractions concepts and procedures</td>
<td>- concrete-representational-abstract (CRA) vs. RA instruction - mastery = 80%</td>
<td>10 lessons/ daily 45 minutes</td>
</tr>
<tr>
<td>Witzel, Mercer, &amp; Miller (2003)</td>
<td>algebra</td>
<td>- CRA sequence of instruction vs. abstract instruction</td>
<td>19 lesson sequence/5-minute class sessions</td>
</tr>
<tr>
<td>Xin, Jitendra, &amp; Deatline-Buchman (2005)</td>
<td>word problems-solving/multiplication &amp; division (multiplicative compare &amp; proportion problems); 1-step problems</td>
<td>- schema-based (SBI) (schema diagram) - general strategy instruct. (GSI/pic.) - mastery = 100% 2 consecutive sessions</td>
<td>SBI: 12 sessions (each type, mixed GSI: 12 sessions (both types each session)</td>
</tr>
<tr>
<td>Reference</td>
<td>Design</td>
<td>Dependent Variables</td>
<td>Posttest /ES²</td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>---------------------------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Bottge, Rueda, Serlin, Hung, &amp; Kwon (2007)</td>
<td>nonequivalent dependent variables design</td>
<td>- KKC: rates, interpret data, construct graphs, make predictions&lt;br&gt;- FCC: length, whole # combinations, fractions, interpret data</td>
<td>•Kim’s Komet Challenge&gt; Fraction of the Cost Challenge: 1.44&lt;br&gt;•Students with LD in inclusive classes scored slightly lower than classmates pretest/posttest</td>
</tr>
<tr>
<td>Butler, Miller, Crehan, Babbit, &amp; Pierce (2003)</td>
<td>pretest-posttest; random assignment of class to teacher</td>
<td>- ratio &amp; proportion problems&lt;br&gt;- part/whole disc. (write fractions for parts)&lt;br&gt;- equivalent fraction&lt;br&gt;- word problem solving: 12-abstract fraction problems</td>
<td>•CRA &gt; RA: 0.16</td>
</tr>
<tr>
<td>Witzel, Mercer, &amp; Miller (2003)</td>
<td>pre-post-follow-up design; random assign. by classroom, matched pairs</td>
<td>- 5 algebra constructs: reducing expressions, inverse oper. negative/divisor var., transformations one side of and across equal sign</td>
<td>•CRA &gt; Control: 0.71</td>
</tr>
<tr>
<td>Xin, Jitendra, &amp; Deatline-Buchman (2005)</td>
<td>pretest-posttest comparison group design; random assign. to groups</td>
<td>- word problem solving: 16 one-step multiplication and division word problems</td>
<td>SBI &gt; GSI: +1.69</td>
</tr>
</tbody>
</table>
Results: Question #2

• Domain

  • Of the 9 studies,
    • 7 (#1, 2, 4, 5, 6, 7, 9) focused on word problem solving: 1 study included measurement, 4 of the studies involved computation, 2 of the studies taught fractions, and 1 study incorporated measurement, computation, fractions, decimals, percent, algebra, and data analysis and probability (#6)
    • 1 (#8) study focused on just teaching algebra
    • 1 (#3) study focused on just teaching fractions

• Intervention

  • Of the 9 studies,
    • 4 of the 5 single-subject studies (#1, 2, 4, 5) included maintenance and generalization conditions; for the group studies, 1 study (#9) showed a large effect size for maintenance and medium for generalization; findings were mixed or the condition was not included for the remaining 3 studies (#6, 7, 8)
Results: Question #2

- Effectiveness of intervention

  - Of the 5 single-subject studies,
    - 3 of the studies (1, 2, 3) demonstrated PND greater than 69%, which is considered “effective;” schema-based instruction, strategy instruction, or use of manipulatives were interventions employed in these 3 studies; 1 study (# 4) showed slightly below (66.6%) “effective” for teaching representation and effective (94.3 %) for solution strategies; 1 study (# 5) had results in the “questionable” range; the intervention was schematic diagrams; maintenance and generalization results were mixed

  - Of the 4 group studies,
    - 2 studies (# 6, 9) had large effect sizes employing anchored instruction (#6) or schema-based instruction (#9); 1 study (# 8) had a medium effect size and included a CRA routine; and 1 study (# 7) demonstrated a small effect size and also employed a CRA routine
Results: Question #3

**Single Subject Studies: n = 5**

- **Description of Participants and Setting**
  - 100% of the studies (n = 5) participants were described with sufficient detail including gender and ethnicity
  - 100% of the studies included LD status information
  - The majority (80%) of the studies (# 1, 2, 4, 5) provided sufficient detail on participant selection and 2 studies (40%) (# 1, 2, 4) offered detail on the physical setting

- **Dependent Variable**
  - 100% of the studies described the dependent variable with operational precision and procedures were used to yield a quantifiable index
  - In all of the studies, measurement of the dependent variable was valid and described with replicable precision and the dependent variables were measured repeatedly over time
  - Reliability data were collected and IOA standards (80% or higher) were achieved in all of the studies
Results: Question #3

Single Subject Studies: Continued

• Independent Variable
  • 80% of the studies (# 1, 2, 4, 5) described the independent variable with replicable precision
  • Systematic manipulation of the IV was achieved in all of the studies; however, fidelity of implementation was noted in only 60% (n = 3) of the studies (# 1, 2, 4)

• Baseline and Experimental Control/Internal Validity
  • Repeated measurement of the DV was evident in all studies; all studies described baseline conditions with replicable precision
  • Experimental effects at three different points in time were documented across all of the studies yet threats to internal validity were accounted for in only three of the studies (# 1, 2, 3, 5)
  • In all of the studies, results documented a pattern that showed experimental control

• External and Social Validity
  • Experimental effects were replicated in all studies across participants (multiple baseline/multiple probe)
  • The DV was deemed socially important by the participants (students/teachers) across all of the studies; the magnitude of change in the DV and practicality and cost effectiveness of the IV were also evident in all of the studies; findings were mixed regarding the implementation of the IV across time, interventionists, and settings (# 1, 2, 3 included features that satisfied this indicator)
Results: Question #3

*Single Subject Studies: n = 5*


Results: Question #3

Group Design Studies: n = 4

• Conceptualization Underlying the Study
  • All of the studies made a compelling case for the importance of the research; 1 study (# 6) included research questions stated for the purposes of the study

• Participants/Sampling
  • 2 (50%) of the studies (# 8, 9) employed procedures that were comparable across intervention conditions and 2 (50%) of the studies (# 6, 9) included sufficient information to determine a disability

• Intervention and Nature of Comparison Condition(s)
  • 3 (75%) of the studies (# 6, 7, 9) clearly described the intervention and all of the studies described procedures for assessing fidelity
  • 3 (75%) of the studies (# 6, 7, 8) described instruction in the comparison condition
Results: Question #3

Group Design Studies: Continued

• Outcome Measures
  • 3 (75%) of the studies (# 6, 7, 9) employed multiple measures that aligned with the intervention and measured generalized performance
  • 1 study (# 6) included information about different types of reliability (e.g., internal consistency, inter-rater) and all of the studies included measured the intervention’s effect at appropriate times
  • In only 1 (25%) study (# 9) were data collectors blind to study conditions
  • 2 (50%) of the studies (# 6, 7) provided evidence of the validity of the measures (criterion related and construct validity)
  • 3 (75%) studies (# 6, 8, 9) measured the intervention’s effect beyond the immediate posttest

• Data Analysis
  • 2 (50%) of the studies (# 7, 9) reported Effect Sizes
  • None of the studies reported attrition rates among intervention samples
  • 3 (75%) of the studies (# 7, 8, 9) accounted for variability within the sample
Results: Question #3

Group Design Studies: n = 4


Discussion

Findings

• The majority of the studies identified for this synthesis taught word problem solving using schema-based or strategic-based instruction, anchored instruction, and CRA. Two studies focused on teaching skills and concepts with either CRA or a specific strategy. Mixed effects across studies could be attributable to variations in duration/frequency or to the measure (researcher-developed vs commercially developed).

• Domains included primarily word problem solving; number & operation (fractions, decimals, percent, computation), measurement, algebra, and data analysis & probability, which were included in word problem solving or taught alone. Domain focus in the majority of studies aligns with middle school curriculum in preparation for high school math.
Discussion

Findings

• Inclusion of variables that satisfied the quality indicators recommendations was mixed across the studies. Six of the studies were published before 2005, 1 was published in 2005, and two were published after 2005. Given that the quality indicator recommendations were published in 2005 and this study spanned the years 2000-2008, we would expect inclusion of the indicators to increase in future special education studies.

Limitations

• The main limitation of this study is the limited number of studies that were identified using our selection criteria. Findings should be interpreted cautiously.
Discussion

Future Research

- Future studies should focus on broadening the age/grade range to determine effective interventions in high school. As the curriculum becomes more difficult, it is important for middle school and high school special education teachers to have access to evidence-based interventions to teach complex skills and concepts.

Implications for Classroom Practice

- Implications for classroom instruction include implementing schema-based and strategic-based interventions to teach word problem solving. Previous research findings regarding CRA routines have been effective especially for elementary-age students. This routine should be further explored for older students.
References


References


