Using Student Work Samples to Promote Calculus Learning

L. Jeneva Moseley, University of Tennessee
Jo Ann Cady, University of Tennessee
Kris Cozad, Greenville County School System
Problem Statement

• NCTM suggests students communicate mathematically by critiquing other students’ work, including incorrect work
  - (NCTM, 2000; Kramarski & Zoldan, 2008; Borasi, 1994).

• Examining incorrect examples may lead to duplication of the errors
  - (Kramarski & Zoldan, 2008; Borasi, 1994).

• Should mathematics teachers have students critique correct or incorrect solutions to improve mathematical proficiency?
Research Questions

If one group analyzes correct work samples and the other group analyzes incorrect work samples, what difference, if any, exists between groups in

• final exam scores?
• determining correct solutions to problems similar to the ones analyzed?
• whether students replicate errors similar to the incorrect work samples?
Research Questions

• What difference, if any, exists between groups in perceptions of how the analyses of student work samples increases mathematical proficiency?

• How do students describe their experiences analyzing work samples?
Theoretical Perspective

The four components of the IMPROVE framework was used to guide our study.

(a) comprehension of the problem,
(b) connections from prior knowledge to new knowledge,
(c) use of appropriate strategies, and
(d) reflection on the process and solution

(Mevarech and Kramarski, 1997)
Theoretical Perspective

Examples of IMPROVE questions:

• How is this similar or different to prior knowledge or a different problem?

• What are the strategies appropriate for solving this, and why?

• What did I do wrong?

• Does this solution make sense?

• How can this be worked a different way?

(Mevarech and Kramarski, 1997)
Review of Literature

- **Attending to errors**
  (Cherepinsky, 2011; Kasman, 2006; Son & Moseley, 2012; Zerr & Zerr, 2011)

- **Writing about mathematics**
  (Green, 2002; Kasman, 2006; Son & Moseley, 2012; Stalder, 2001)

- **Recognizing valid non-traditional solution methods**
  (Blythe, Allen, & Powell, 1999; Driscoll & Moyer, 2001; NCTM 2001; Katims & Tolbert, 1998; Kelemanik, Janssen, Miller & Ransick, 1997; Saxe, Gearhart & Nasir, 2001)
Review of Literature, continued

• Deciphering student work is difficult (Ball, 1990, 1997, 2001; Even & Markovitz, 1995; Even & Tirosh, 1995; Schifter, 2001; Chamberlin, 2005).

• More mathematically in-depth responses can be elicited by incorrect student work samples than by correct ones (Son & Moseley, 2012).

Review of Literature, continued

• Asking students to detect errors of fictional characters, with nonconventional strategies (Kasman, 2006)

• Intervention comparisons:
  – diagnosing errors,
  – self-questioning using IMPROVE questions,
  – combination of diagnosis of errors and self-questioning.

• The combined approach was most effective (Kramarski and Zoldan, 2008).
Participants

• 181 Basic Calculus students

• Two groups: one analyzed correct student work samples and one group analyzed incorrect student work samples.

• Stratified random sampling for the quantitative part

• Purposive sampling for the qualitative part
Treatments

• Learning modules (LMs) created based on identified common errors
• LMs included incorrect and correct samples
• Participants respond to work samples
• LMs give participants examples of how other participants might respond.
• Followed by opportunity to say more
• 5-question survey at the end of each LM
Here is an example of a Learning Module.

https://learningmodules.papyrs.com/Allison
Correct Student Work Sample

Find the derivative of \( f(x) = x^2 + 3x \) using the limit definition.

\[
\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\
\lim_{h \to 0} \frac{[(x+h)^2 + 3(x+h)] - [x^2 + 3x]}{h} \\
\lim_{h \to 0} \frac{x^2 + 2xh + h^2 + 3x + 3h - x^2 - 3x}{h} \\
\lim_{h \to 0} \frac{2xh + h^2 + 3h}{h} = \lim_{h \to 0} h(2x + h - 3) \\
\lim_{h \to 0} 2x + h - 3 = 2x + 0 + 3 \\
\lim_{h \to 0} 2x + 3 = 2x + 3
\]

Briefly describe the steps Melissa took. Do you think this solution is correct?

From your understanding of what a derivative is, does Melissa’s method make sense? Explain.

Do you think this method will work every time, or could there be special circumstances that would prevent this method from working? Explain.

Can you share another way of finding this derivative?
Here’s an example of one showing incorrect work.
https://learningmodules.papyrs.com/Sharon
Incorrect Student Work Sample

Find the derivative of $f(x) = x^2 - 3x$ using the limit definition.

Limit definition: $f'(x) = \lim_{h \to 0} \frac{f(x+h)-f(x)}{h}
= \lim_{h \to 0} \frac{(x^2+3x+h)-(x^2+3x)}{h}
= \lim_{h \to 0} \frac{x^2+3x+h-x^2-3x}{h}
= \lim_{h \to 0} \frac{h}{h}
= 1 = \boxed{1}$

Uh oh! The answer should be $2x+3$.

Briefly describe the steps Lisa took. Why do you think this solution is incorrect?

What would you say to Lisa to help with the problem?

From your understanding of what a derivative is, does Lisa’s answer seem reasonable or unreasonable? Explain.

What specific steps or strategies could Lisa use to avoid this type of error?
Find the derivative of \( f(x) = x^3 + 3x \) using the limit definition:

\[
\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}
\]

Limit definition: \( f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \)
Five Question Survey

• Purpose is to determine students’ perceptions regarding examining student work samples

• Framework to guide their reflections
  – Mathematical proficiency (NRC, 2001)
    • conceptual understanding
    • procedural fluency
    • strategic competence
    • adaptive reasoning
    • productive disposition (p. 5).
Five Question Survey

Using a five-point Likert Scale, rate the helpfulness of the learning module as it related to your

• understanding of mathematical ideas.
• ability to follow rules of math.
• ability to develop a problem-solving strategy.
• ability to think about and explain your answers.
• attitudes about math about your abilities.
Learning Modules 7 & 8

• Two work samples were provided instead of 1
• Incorrect modules composed of two different errors
• Correct modules composed of traditional vs. nontraditional solution strategy
• Questions include “Which is better or worse?”
• Followed by same five-question survey
Data Sources

- Pretest
- Semester Final exam
- Four tests throughout the semester
- Survey to determine perceptions of helpfulness
- Interview
Data Analysis

• Analysis of Covariance (ANCOVA) was used to analyze (1) final exam scores, (2) points earned on certain test questions that are similar to those in LMs, (3) percentage of instances in which similar errors were made, while taking pretest data into consideration.

• Independent samples t-test was used to analyze the Likert-scale rating means.

• For the qualitative data, I used an inductive analysis (Hatch, 2002).
Results

• Results of the ANCOVA revealed no significant differences ($p = 0.267$, $F = 1.239$, $df = 1$) in final exam scores between the group that analyzed correct work samples (adjusted mean = 80.782) versus the group that analyzed incorrect student work samples (adjusted mean = 82.788), after accounting for the pretest scores.
Results

- Results of the ANCOVA revealed no significant differences ($p = 0.367$, $F = 0.819$, $df = 1$) in the correctness of problems similar to those in learning modules, measured by an overall percentage of points earned on those problems, between the group who analyzed correct work samples (adjusted proportion = 78.342%) and the group who analyzed incorrect work samples (adjusted proportion = 79.920%), accounting for pretest points earned on similar problems.
<table>
<thead>
<tr>
<th>LM</th>
<th>Adjusted Means</th>
<th>ANCOVA stats</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct</td>
<td>Incorrect</td>
</tr>
<tr>
<td>LM1</td>
<td>94.683</td>
<td>91.872</td>
</tr>
<tr>
<td>LM2</td>
<td>86.066</td>
<td>86.350</td>
</tr>
<tr>
<td>LM3</td>
<td>85.055</td>
<td>87.263</td>
</tr>
<tr>
<td>LM4</td>
<td>76.401</td>
<td>73.853</td>
</tr>
<tr>
<td>LM5</td>
<td>84.462</td>
<td>84.716</td>
</tr>
<tr>
<td>LM6</td>
<td>87.186</td>
<td>90.563</td>
</tr>
<tr>
<td>LM7</td>
<td>61.342</td>
<td>62.633</td>
</tr>
<tr>
<td>LM8</td>
<td>45.299</td>
<td>42.150</td>
</tr>
</tbody>
</table>
Results

Results of the ANCOVA revealed no significant differences \((p = 0.123, F = 2.406, df = 1)\) in the replication of errors similar to those in learning modules, measured by a percentage of times the error was made given the opportunity to make the error, between the group who analyzed correct work samples \((\text{adjusted proportion} = 8.3\%)\) and the group who analyzed incorrect work samples \((\text{adjusted proportion} = 7.3\%)\), accounting for frequency of the errors made on the pretest.
<table>
<thead>
<tr>
<th>Error Type</th>
<th>Adjusted Means</th>
<th>ANCOVA stats</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct</td>
<td>Incorrect</td>
</tr>
<tr>
<td>Illegal Cancel</td>
<td>0.025</td>
<td>0.033</td>
</tr>
<tr>
<td>f(x) + h Error</td>
<td>0.053</td>
<td>0.018</td>
</tr>
<tr>
<td>Plus or Minus</td>
<td>0.083</td>
<td>0.062</td>
</tr>
<tr>
<td>Exponent Rule</td>
<td>0.085</td>
<td>0.070</td>
</tr>
<tr>
<td>Quotient Rule</td>
<td>0.050</td>
<td>0.022</td>
</tr>
<tr>
<td>Ordered Pair</td>
<td>0.079</td>
<td>0.043</td>
</tr>
<tr>
<td>Quotient Integration</td>
<td>0.013</td>
<td>0.10</td>
</tr>
<tr>
<td>Plus C Error</td>
<td>0.344</td>
<td>0.478</td>
</tr>
<tr>
<td>Power Rule on Exponential</td>
<td>0.042</td>
<td>0.082</td>
</tr>
</tbody>
</table>
Results

Results of the independent samples t-test revealed significant differences \((p = 0.009, t = 2.632, df = 1142)\) in perception of helpfulness to mathematical proficiency, measured by the mean of the means of the Likert scores, between the group who analyzed correct work samples \((mean = 3.3935, std. deviation = 1.16809)\) and the group who analyzed incorrect work samples \((mean = 3.2060, std. deviation = 1.24015)\).
Results

When considering all survey results collected across time, over the entire semester, statistically significant group differences were found in perceptions of helpfulness to each strand of mathematical proficiency except for productive disposition.
<table>
<thead>
<tr>
<th>Strand</th>
<th>t</th>
<th>Sig.</th>
<th>Group Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptual Understanding</td>
<td>t=2.752</td>
<td>P=0.006</td>
<td>3.36</td>
</tr>
<tr>
<td>Procedural Fluency</td>
<td>t=2.432</td>
<td>P=0.015</td>
<td>3.41</td>
</tr>
<tr>
<td>Strategic Competence</td>
<td>t=2.609</td>
<td>P=0.009</td>
<td>3.40</td>
</tr>
<tr>
<td>Adaptive Reasoning</td>
<td>t=2.561</td>
<td>P=0.011</td>
<td>3.55</td>
</tr>
<tr>
<td>Productive Disposition</td>
<td>t=1.857</td>
<td>P=0.064</td>
<td>3.24</td>
</tr>
</tbody>
</table>
Results

LMs improved **conceptual understanding** because

- they made students think.
- they asked why.
- they prompted students to seek out information.
Results

LMs positively influenced **procedural fluency** because:

- they made students think.
- students had to seek outside information/help.
- they highlighted dilemmas of how to use rules.
- students explained procedures and gave advice.
- students worked the problems independently.
- students made connections between steps.
- avatars gave immediate feedback.
Results

LMs positively influenced *strategic competence* because:

- they made students think.
- they asked why.
- they showed how to begin problem-solving.
- students worked the problems independently.
- they encouraged students not to skip steps.
- they highlighted dilemmas of how to use rules.
- they were open-ended, with choices, instead of right or wrong.
Results

A factor that negatively influenced strategic competence was predictability (in group who always analyzed correct work samples) in that the expectation that the work was correct did not make students think as much.
Results

LMs improved **adaptive reasoning** because

• they made students think.
• they asked why.
• they made students put it into words.
• they created a “safe” environment to express ideas without embarrassment.
Results
LMs improved **productive disposition** because
• lack of pressure to submit perfect answers.
• they showed students that mistakes happen and are fixable (among those who analyzed incorrect work samples).
• content aligned with class and was accessible.
• avatars gave immediate feedback.
• they encouraged and motivated.
• they helped to improve students’ feelings about their own math abilities.
Results

Factors that negatively influenced productive disposition were:

• the feeling that the LMs were extra work.
• a feeling of futility.
• predictability (among those who analyzed correct work samples).
Discussion

• Like NCTM suggests (2000), students should communicate mathematically by critiquing student work samples.

• Teachers’ and researchers’ assumptions that examining incorrect examples may lead to duplication of errors are not confirmed by this study.
Discussion

• We knew that deciphering student thinking was difficult for teachers (Chamberlin, 2005), and this study shows that requiring math students to decipher student thinking stretches them mathematically, while only having limited negative influences on attitudes.
• Son and Moseley (2012) found more mathematically in-depth responses to be elicited from incorrect work samples; however, this study did not uncover significant differences in other variables, such as replication of errors, problem-solving abilities, and overall achievement.
Discussion

• Although Son and Moseley (2012) found more in-depth responses to be elicited from incorrect work, this study found that correct work samples are perceived by students to be more helpful at the time of analysis.

• This difference may be due to sample size.

• Some counterevidence was uncovered in the interviews.
Discussion

• As with Kasman’s study (2006), use of fictional characters created a “safe” environment for conversing about math.

• Kramarski and Zoldan (2008) found a combination of self-questioning and error-diagnosis to be most effective, but the interviews in this study revealed more information about how and why the use of this combination influences learning.
Conclusions

• Because there were not any significant differences in replication of errors, problem-solving, or overall math achievement, teachers should not avoid using incorrect work samples, provided that students are thinking and analyzing the errors.

• Analyzing correct work was perceived to be more helpful at the time of analysis, but interviews uncovered how both types improved mathematical proficiency.
References


References


References


References


References


References


References

