Game Development and Design: Blood, Sweat, and Tears

Greg Chung

Warp Speed, Mr. Sulu: Integrating Games, Technology, and Assessment to Accelerate Learning in the 21st Century

Redondo Beach, CA – April 29-30, 2014
Structure of Talk

- Blood: Large-scale randomized controlled trial (RCT)
- Sweat: How did we get there?
- Tears: Innovations and impact
Project Goal

- Develop and evaluate video games aimed at improving students’ knowledge and skills in pre-algebra topics
  - Focus on rational numbers, solving equations, functions (middle school)
  - Develop and test games
  - RCT: Examine effects of game on learning of fraction concepts
The Blood: Randomized Controlled Trial
Research Question

• Does playing CATS-developed rational number games result in more learning of rational number concepts compared to playing a comparison set of games?
Sample

- 62 classrooms randomly assigned to condition
  - 30 treatment classrooms
  - 29 control classrooms
  - 3 classes dropped because of schools’ technology
  - ~1700 students

- Final sample—59 classrooms, 9 districts, 24 schools, CA and NV, ethnically diverse
Procedure

• Teachers integrated math games into their curriculum
  ✓ Rational number or control topic
  ✓ 3 hours professional development
  ✓ Implementation window: Jan to April 2012 (pre-state testing)
    ❖ Pretest (30min)
    ❖ 10 gameplay days within window (40min per occasion)
    ❖ Delayed posttest 1 week after last game (30min)

• $600 honorarium to teacher
Measures

3. The figure below shows $\frac{3}{7}$ of a whole unit shaded. Complete the figure to show where the whole unit ends. Be sure to draw lines ("|") to show where each piece is.

8. At what number is the "?" located?

13. $\frac{1}{6} + \frac{1}{4} = \frac{\square}{\square}$

18. $\frac{8}{9} \times \frac{\square}{\square} = 1$
Results

- The treatment condition performed significantly higher on the posttest than the comparison condition
  - Treatment condition \((n = 30)\) played 4 games related to fractions
  - Comparison condition \((n = 29)\) played 4 alternative games (solving equations)
  - No differences between conditions on pretest
Students were engaged in almost every game. At least two games helped students learn.

Would use games again in classroom

Teacher’s Perceptions

At least two games helped students learn
Students were engaged in almost every game
Would use games again in classroom

OMG!!

OMG!!
Efficacy Trial Results

- Multi-level model shows significant effects of treatment, effect size = 0.6 (Li Cai discussing next session)

<table>
<thead>
<tr>
<th>Type of Intervention</th>
<th>Mean Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructional format</td>
<td>0.21&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Instructional component/skill training</td>
<td>0.36&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Focused on individual students</td>
<td>0.40&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Simulation and games</td>
<td>0.34&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Computer-assisted instruction</td>
<td>0.31&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>


The Sweat: How Did We Get There?
RCT Launch Schedule

RCT study lock
• Identified final set of games, measures, PD design

Procedure test
• Test PD, gameplay, administration procedures
• Tune PD, games, measures

Full dry run
• 1 classroom

Launch

Game Development Studies

Y1

Y2

Y3

Y4

• 40 CRESST Laptop usage
  • Feb: 100% of days
  • Mar: 76% of days
  • Apr: 23% of days

• Data collection activity
  • Feb: 100% of days
  • Mar: 100% of days
  • Apr: 95% of days
Game Development Studies
2 Years Before RCT

- Total of 24 trials, Grades 4-12
  - Setting: Schools, after-school program, private school
    - 18 sites; 8 school districts; 4 states

- Types of studies
  - Small scale tryouts: pre-post, think-alouds
    - 7 Studies; total N=238
  - Experimental studies: Instructional variations
    - 10 studies; total N=1588
    - Typically 30-40 minutes
  - Assessment and methodology studies
    - 7 studies; experimental data plus 711 new subjects
Game Development and Testing Process

1. **Development**
   - Establish instructional sequence + assessments

2. **Game Prototype + Testing**
   - Small-scale (1-5 students)
   - Classroom-scale (30 students)

3. **Play/Paper Testing**
   - Small-scale (1-5 students)
   - Classroom-scale (30 students)

4. **Game Effectiveness Trials**
   - 60 + kids per condition

5. **Large Scale Efficacy Test**
   - 59 classrooms (1,700 kids)
Fractions Games

- **WikiJones** (number line)
- **Save Patch** (unit, fractional pieces, adding fractions)
- **Tlaloc’s Book** (inverse operations)
- **Rosie’s Rates** (functions)
Solving Equations Games

**MonsterLine**
(operations on pos. and neg. integers)

**Expressso**
(transforming expressions)

**Zooples**
(solving equations)

**AlgebRock**
(solving equations)
Game Development Studies

Results

• Empirical

✓ Instruction/feedback – less is more, narrative and avatar choice matters for engagement, collaboration helpful for low performers, repeated play improves learning, math measure validation, engagement scale validation, misconceptions and play strategies detectable via data mining, ...

• Student and teacher self-reports, classroom observations

✓ Game features – music, graphics, chevos, avatar select

✓ Math not really an issue, high engagement, surprises

✓ Technology issues
Game Development Studies
Results

• Developing games for learning -- hard but doable
  ✓ Big difference between teaching and practice
  ✓ Games don’t have to be AAA quality
    ✷ Classroom context sets a low bar
  ✓ Stickiness possible -- the more challenging, the more sticky
  ✓ Game mechanics that require use of knowledge is a key design feature
  ✓ Anecdotal collateral benefits of gameplay -- more student engagement, more individual attention, better feedback to teacher, increased confidence of student, productive student-student interactions
The Tears: Impact and Innovations
Impact: Developed learning effective games

- Kids learned something from our fractions games as measured by an external *transfer* test
Other Impacts

- New methodology
  - See Cai presentation, 10:45-12:15pm session

- Technical reports craziness
  - CATS technical reports downloaded 69,456 times

- Journals (15), proceedings (7), chapters (9), technical reports (22), conference presentations (93)

- Transfer of CATS-based approaches and methodologies to other game-focused projects at CRESST (DARPA, ONR, PBS KIDS, CSUSB, ...)

Innovation: Coherent Design Process

- Ontologies and knowledge specifications
- Game design, assessment, professional development
- Guides what to teach in game, what to measure, what to focus on in professional development
Innovation: Game Testbed to Accelerate R&D

<root>
    <verifylogins> false </verifylogins>
    <loginfile> RNLogins.xml </loginfile>
    <hidescroll> false </hidescroll>
    <cheats> true </cheats>
    <newFormat> true </newFormat>
    <showbuild> true </showbuild>
    <lockcharacters> true </lockcharacters>
    <levelselect> true </levelselect>
    <onlyvisited> false </onlyvisited>
</root>

<LevelName> Level 50 </LevelName>
    <LevelWidth> 3 </LevelWidth>
    <LevelHeight> 2 </LevelHeight>
    <LevelDenominator> 2 </LevelDenominator>
    <LevelScale> 2 </LevelScale>
    <NumKeys> 0 </NumKeys>
    <NumCoins> 0 </NumCoins>

<Sign>
    <xNum> 0 </xNum>
    <xDenom> 2 </xDenom>
    <yNum> 4 </yNum>
    <yDenom> 2 </yDenom>
    <required> false </required>
    <optional> false </optional>
    <goal> false </goal>
    <start> true </start>
</Sign>
Innovation: Gameplay as a Data Source

- Key design properties

  ✓ Game telemetry reflects description (not inferences) about events connected to learning

  ✓ Game telemetry packaged in a structured format (i.e., usable)

  ✓ Game mechanic requires use of math knowledge

  ✓ Game allows player to fail (i.e., commit errors)

  ✓ Game requires player to make decisions (do I do this or that?)

  ✓ Player cannot progress without requisite knowledge

  ✓ Stage / level design useful [stages = 1 concept; levels = variations of concepts]; sawtooth curve
Dear UCLA, ever since 1st grade (I'm not lying) I've never been good in math. But I'm not stupid, not at all. Your games have taught me a lot. I want to thank all you guys for helping me out with my struggle with math. Many thanks.