Overview

• Objective
• Design History & Issues
• Next Steps
Design History

• Design Review of hands-on version
• Identify Key Design Issues for virtual version
• Literature Review
• Scenarios & Storyboards (paper based)
• Interface and interactions requirements
• Flash storyboard
• Flash mockup
• Initial user testing (formative)
Paper Towels-The Hands On

- Process Open/Content Lean
- Interater Reliability: Very High (.94) – Training Can Produce Consistent Results.
- Intertask Reliability: Low – Many Tasks Needed to Generalize Performance to the Domain.
- Sources of Error: Very Few (i.e. Saturation, Care?)
Instructional Goals

• Assessment and development of scientific reasoning, critical thinking, and inquiry.
• Students design and conduct sound scientific experiments, including identifying and controlling variables.
• Students can apply previously learned skills.
• Develop appreciation and skills required for careful observation and measurement.
## Conceptual Analysis

<table>
<thead>
<tr>
<th></th>
<th>Declarative</th>
<th>Procedural</th>
<th>Schematic</th>
<th>Strategic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Performance Task</strong></td>
<td>What is… weight, volume &amp; relationship between them saturation area &amp; length, purpose/use of different tools</td>
<td>How to… read/write saturate control variables measure volume, weight etc. determine amount of water soaked up</td>
<td>Why… saturate control variables use care in procedures</td>
<td>recognize when has reached valid/reliable result recognize when unexpected factors affect result</td>
</tr>
<tr>
<td><strong>Response Format</strong></td>
<td>Which towel absorbs the most and which the least?</td>
<td>Steps in procedure</td>
<td>How did you know? Completely wet? Same size?</td>
<td></td>
</tr>
<tr>
<td><strong>Scoring System</strong></td>
<td>Correct result</td>
<td>Method for getting towel wet Determine result Care in procedure</td>
<td>Saturation Determine result Care in procedure</td>
<td>Care in procedure</td>
</tr>
</tbody>
</table>
What is the utility of this hands-on performance assessment?

(+)  
- Provides richer feedback regarding the student’s ability to solve problems using scientific reasoning  
- The experience is instructional and non-intimidating  
- Allows you to characterize students’ procedural strengths and weaknesses (good feedback to teacher)  
- It’s much more collaborative and fun!

(-)  
- More costly than multiple-choice, paper-pencil test (new and additional hands-on materials)  
- Requires more time (cost) to score and be trained to score; requires trained scorers (harder to score)  
- Requires more set-up time  
- Requires more class time to execute the test  
- The lab-book depends on literacy skills  
- It’s weak in content
Implications for Virtualizing PT

• Real world mistakes vs. virtual world mistakes - How many can/should we accommodate?
• How many roads (scenarios) can we account for?
• Cross-method results (combining results from multiple experiments)
• Can human observation really be replaced by computer tracking?
• No notebooks, so should we bring back “care”
• Do we still need to give them some hands-on materials?
• Can we really do this??
## Conceptual Analysis

<table>
<thead>
<tr>
<th>Performance Task</th>
<th>Declarative</th>
<th>Procedural</th>
<th>Schematic</th>
<th>Strategic</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is… weight, volume &amp; relationship between them saturation area &amp; length, purpose/use of different tools</td>
<td>How to… read/write saturate control variables measure volume, weight etc. determine amount of water soaked up</td>
<td>Why… saturate control variables use care in procedures</td>
<td>recognize when has reached valid/reliable result recognize when unexpected factors affect result</td>
<td></td>
</tr>
</tbody>
</table>

| Response Format | Which towel absorbs the most and which the least? | Steps in procedure | How did you know? Completely wet? Same size? | |

| Scoring System | Correct result | Method for getting towel wet Determine result Care in procedure | Saturation Determine result Care in procedure | Care in procedure |
Key Design Issues

• Content issues
  - Should the notebook be included in the computer-based version? Should the students have to take notes?
  - Should all the items present in the hands-on version also be included in the computer-based one?

• Design decision:
  – computer-based should be faithful to hands-on
  – include the notebook and all items
Key Design Issues, cont'd

• Interface issues
  – Support all possible interactions? (wiping the scissors, dunking a roll, etc.)
  - Separate working area where interactions can take place?
  - 2D or 3D?
  - Where should the notebook be placed?

• Design decision:
  – Support only interactions necessary for supported.
  – 3D perspective is employed
  – separation between the working area and storage
  – working area should be maximized
  – notebook placement TBD through user testing
Key Design Issues, cont d

• Scoring issues
  – Should computer or human do scoring?
  – Should care be part of grading?

• Design decisions:
  – Care not tracked, not be a factor in grading
Literature Review

• Authentic assessment and learning and knowing:

• Exchangeability
  – Rosenquist.

• Performance assessments and science:

• Paper Towel Assessment:

• Using technology for assessments:
  – Bennet, 1999
Scenarios and Storyboards: 
1) Weight of Towels

1. materials to be used: scale, paper towels, pitcher of water
2. pull 1 sheet
3. dip sheet in pitcher of water
4. pull sheet out, let drip
5. weigh sheet and record
6. repeat steps 2-5 for other 2 paper towels
7. compare weights of 3 towels (more weight -> higher saturation)
Scenarios and Storyboards:
2) Amount squeezed out

1. Materials to be used: paper towels, pitcher of water, 3 beakers, 1 tray
2. Pull 1 sheet
3. Dip sheet in pitcher of water
4. Pull sheet out, let drip into tray
5. Squeeze PT and tray water into empty beakers
6. Repeat steps 2-5 for other 2 paper towels
7. Compare by level of water in beaker (More water -> higher saturation)
Scenarios and Storyboards: 3) Remaining beaker water

1. materials to be used: paper towels, pitcher of water, 3 beakers, trash can
2. fill beakers with equal amounts of water
3. put towel 1,2,3 in corresponding beaker
4. observe saturation
5. remove PT from each beaker & let free drip into beaker
6. discard paper towel
7. compare by level of water in beaker (less water-> higher saturation)
Storyboards for Computer simulation (example)

Goals for activity scenarios

- identify the characteristics of target audience, background and expectations
- discover details of activities they go through, problems encountered, thought processes
- identify the key features affecting users
- raise more usability questions to pursue
- enable PT team to decide on key issues
Storyboard Actor

Johnnie is a 5th grader at Palo Alto elementary school
•  His grade for Science is above average.
•  Science has been his favorite subject.
•  He likes to do Science experiment since he thinks it is fun.
•  His teacher is preparing a paper towel computer simulation for his class today.
•  Students will go to the computer lab and do the simulated experiment individually.
Activity Scenario—“Squeeze water out approach (I)

Step 1: Approach workbench, read instructions, and planning
¥ 1. Johnnie sees tools and objects neatly displayed on the shelf
¥ 2. workbench with a notebook with instructions
¥ 3. introductions to materials
¥ 4. Johnnie’s approach to the PT is see how much water was in the paper towels after they have been dunked by squeezing out the excess.

°Step 2. Johnnie tears off a piece of PT A
(roll over message says click to tear off a piece of PT A).

°Step 3: Dip PT A into a pitcher of water
°Step 4: Pulled out the wet PT and let it dripped into tray
Activity Scenario—“Squeeze water out approach (II)

Step 5: Squeeze the wet PT into beaker A and record the height.
   He recorded his reading on the electronic notebook.

Step 6: Repeat step 2-5 for rolls B and C
   The sizes of the PT from roll A, B or C were the same.

Step 7: Compared the heights and conclude his answers, using his electronic notebook

Step 8: Repeat 2-7 for confirmation--- just to be on the safe side!
   A dialogue pops up do you want to keep your previous record?
   Yes/no . Johnnie clicked yes. His records from the first experiment were kept.

Step 9: Complete the experiment
   When he completed his notebook, he submitted it to Mrs. Smith.
   Johnnie clicked submit on the screen and the result was sent automatically.
The Group Brainstorming
Mockup History 3

Find out which paper towel can hold, soak up or absorb the most/least water.

To begin, just roll your mouse over the objects on the screen. To use them, drag them onto other objects.

Click ‘Go’ to begin.
replace this with your name
Mockup History 5

RESULTS: When you think you know which paper towel can hold, soak, or absorb the most water and the least water, write "most" and "least" beside the name of the towel.

Yellow  Blue  Red
Notebook
replace this with your name
Paper Towels

You have three different kinds of paper towels in front of you and some equipment for doing scientific experiments.

Problems:
1) Find out which paper towel can hold, soak up, or absorb the most water.

2) Find out which paper towel can hold, soak up, or absorb the least water.

Look at each piece of equipment. Think about how you might use some of it to do an experiment to solve the problems. You don't need to use all the equipment.

When you are finished you will be asked to write what you did so one of your friends can repeat the experiment exactly as you did it. You may want to keep notes on a sheet of paper to help you remember what you did and what you found out.
RESULTS: When you think you know which paper towel can hold, soak, or absorb the most water and the least water, write "most" and "least" beside the name of the towel.

Yellow  Blue  Red
Formative Evaluation, Initial User testing

- How do users interact with the PT mockup (B-1 version)?
- How do users perform the procedures they would want to with the real hands-on PT experiment using computer simulation?
Formative Evaluation, Methods

Co-discovery and talk aloud:
- Users were paired up except Subject 5.
- Each pair was taken to a computer on which our prototype had been previously loaded.
- The tester asked the subjects a few questions based on a questionnaire and wrote down their answers.
- The subjects were asked to perform experiment by reading the instructions and then moving the icons while explaining what they are trying to do.
- After all had finished a debriefing session was held.
# Formative Evaluation, Subjects

<table>
<thead>
<tr>
<th></th>
<th>Pair-up</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subject1</td>
<td>Subject2</td>
<td>Subject3</td>
<td>Subject4</td>
<td>Subject5</td>
</tr>
<tr>
<td><strong>Grade (Age)</strong></td>
<td>8\textsuperscript{th} (13)</td>
<td>8\textsuperscript{th} (13)</td>
<td>5\textsuperscript{th} (10)</td>
<td>5\textsuperscript{th} (11)</td>
<td>5\textsuperscript{th} (11)</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td><strong>Good at Science</strong></td>
<td>Good</td>
<td>Good</td>
<td>OK</td>
<td>Ok</td>
<td>Good</td>
</tr>
<tr>
<td><strong>Computer experience</strong></td>
<td>A lot</td>
<td>A lot</td>
<td>A lot</td>
<td>A lot</td>
<td>A lot</td>
</tr>
<tr>
<td><strong>Computer Games experience</strong></td>
<td>A lot</td>
<td>A lot</td>
<td>A lot</td>
<td>A lot</td>
<td>A lot</td>
</tr>
<tr>
<td><strong>Performed PT before?</strong></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

4 December, 2001

Paper Towels Development
Formative Evaluation, Findings

The Notebook:
¥ The functions of +, - keys and draggability not obvious to users
¥ The word next is misunderstood as clickable
¥ A lot of lagging in typing on the notebook as well as dragging it.

The displayed PT elements
¥ Can recognize most of the elements except the trays trays doesn't look like trays .
¥ Paper towel rolls don't look 3D, inconsistent with the rest.
¥ Subject 3 thought the blue PT is soaked (blue means wet to her) and she interpreted the shading effect of the PT as different degree of dampness.
Formative Evaluation, Findings

The interactions
¥ Clicking and dragging the elements are natural to the subjects.
¥ The elements disappear when placing on top of a PT is confusing
¥ When trying to clean the table for repeating the experiment, subject 3 found it troublesome that she needs to drag the element one by one back to the shelf
¥ Confusions mainly arise from the lacking of interactions in our mockup.
Formative Evaluation, User Recommendations

- Give instruction page before simulation begins, explaining that you can click n things and drag them, etc.
- Rollover pop-ups (AKA tool-tips) for all buttons, icons and items.
- Picture or icons on toolbar rather than small items on shelf. Click on one and 3-d object pops up on table. (a sketch representation by subject 5)
- Could have different angles of view of experiment.
Formative Evaluation, User Recommendations, cont'd

- Ways to get things to interact:
  - right click in item to pop-up menu. Menu should have options like Dunk in water for PT sheet, or Pick up water for eye dropper.
  - click on first item then drag a line to the second item
- Drag and drop is not recommended since too much lag!
- Not allow to put things over other things and cover (hide) them.
- Prefer real hands-on experience since it gives you more freedom. It can t/shouldn t replace that with computer based version.
- Computer based should be used for experiments that you can t do in real life (too expensive, too dangerous, etc.)
Formative Evaluation,
Sample sketch

A sketch
by subject 5
Formative Evaluation, Methods employed

**Subject 1:** Poured water into beaker, dunked paper towel, repeat. Compared what was left in the beakers

**Subject 2:** Cut paper towels into smaller pieces and compared how much each paper could hold by squeezing eye-dropper.

**Subject 5:** She will take the blue PT, measure some water into an empty beaker and pour the water onto the blue PT. Then repeat the same procedure for all the 3 PTs. She will decide which holds up more water by looking at the darkness (different shade) of the PT. She will also use ruler to measure the area being soaked. The one with the largest area holds up the most water. To repeat the experiment for confirmation, she will keep all the equipment back to the shelf. Then re-do the experiment again. If she gets the same conclusion, she thinks she is probably right.
Subject 3:
Method 1 -- He will cut PT sheets into 1/8 with scissors, then take eye dropper and fill with water. Counts drops for each piece of PT until water shows up on table around piece.
Method 2 -- Take 3 metal tubs (trays) and pour water into them. Weigh each with scale to make sure they all weigh the same, i.e. have same amount of water. Cut PT sheets into size of trays. Dunk the sheets and let them sit to soak up water. Take out the towels. Weight the trays again to see which has the least amount of water left in it.
Method 3 -- Put sheets on table. Dump water on the sheets. Pick up each sheet with tweezers and weigh with scale to see which is heaviest, i.e. holds most water. Not a very accurate admitted User 1.
Formative Evaluation,
Methods employed, cont d

**Subject 4:**
Method 1: Take vials (meaning graduated cylinders or beakers) and put them on desk. Pour certain amount (say 1 cup) of water into each one from pitcher (same for each). Fold up the PT sheets. Pour water from 1 vial unto each PT sheet. Time how long it takes for water to drop down.

Method 2: Use magnifying class to look at sheet and see which is densest or thickest. Densest holds the most.
Formative Evaluation, Johnnie’s testing

Johnnie tested a separate subject, his son age 10.

¥ good at science
¥ a lot of experience with computers and computer games
¥ had not done the PT experiment before

Method — Talk aloud while using the computer mockup.

Findings — Ben found the icons to be understandable, but wished for greater functionality. It was not clear to him that the notebook could be resized and relocated. He performed some of the experiment on top of the notebook. He found it necessary to clear the table after each step of his experiment which lead to an awkward arrangement of elements. He did not spend time playing with the tools but set about the experiment immediately after reading the directions. He did not make use of the notebook to record notes.
Formative Evaluation, Johnnie's testing, cont d

Procedure for performing Virtual PT — Ben removed single sheets of paper towel and placed them on the table one at a time. He took the eyedropper to the water container. With the idea that the eyedropper contained water, he wet each paper towel. He examined each with the magnifying glass and used the ruler to suggest that he would measure the diameter of the water's spread on each towel.

Recommendation- Ben wanted the objects to be more real in their behaviors. He would like the spread of water to appear as it does with real paper towels.
Summative Evaluation Plan

- Exchangeability: How do we know the computer version works as well as the hands-on assessment?
  - Randomly assign a group to a hands-on version then 2 hours later (approximately) switch that group to do the computer version. Vice versa.
  - Check the correlation between the two assessments. If the correlation is low, perhaps the assessments need to be performed by experts.

\[\begin{align*}
H_1 & \quad H_2 \\
\mid & \quad \mid \\
C_1 & \quad C_2 \\
\end{align*}\]

- \(H_1\) = Hands-On, 1st time
- \(H_2\) = Hands-On, 2nd time
- \(C_1\) = Computer, 1st time
- \(C_2\) = Computer, 2nd time
Summative Evaluation Plan

• Utility
  – For what purposes will this computer version be most useful?
  – What resources will be needed to run this program?
  – How will data be stored and analyzed?
  – How will instruction be modified based on this implementation?
Flash Design Next Steps

- Incorporate User Feedback in redesign of the notebook UI
- Change framerate to 30fps
- Program object interactions and pop-up menu choices
- Add scoring flags and variables that are triggered during interactions
- Add remaining “worksheet” pages to notebook
- Create flash-to-database linkages
- Build database to record online responses
- Create login and registration process
- Analyze the business model and design appropriate accommodations
- Gather stakeholder feedback and do a lot more user testing!
Reflections?