How We Improved Success Rates in Large General Chemistry Classes at the University of Utah

Charles H. Atwood
Brock Casselman
Braden Ohlsen
Chemistry Department
University of Utah
Salt Lake City, UT
Presentation Outline

• Historical Perspective
• Implementation of Required Discussions
• Future Directions
Fail, and You Likely will Never Pass Chemistry!

Historic 21.7% Fail Rate (2000-2012)
61.2% Never Retake the Class
13.4% Retake and Fail
Only 25.4% of All Students who Fail will Ever Pass the Course
Pre-Requisite Implementation

- Correlation Between Math ACT and General Chemistry Performance*
  - Described as Early as 1973: Neil R Coley
  - Predicting Success in General Chemistry in a Community College
  - Math ACT vs College Chemistry Success: $R^2: 0.227$

From OBIA: ACT Scores vs Chem 1210 Pass Rate

\[ y = 0.0173x + 0.3496 \]

\[ R^2 = 0.9364 \]

Pass Rate

ACT Math Score vs 1210 Percent C- or Better

Greater than 25 on Math ACT
Combined Pass Rate: 89.6%

Similar Trend with Math SAT Scores \( (R^2 = 0.759) \)

Pre-Requisites:
25 on Math ACT
600 on Math SAT
Math Accuplacer Equivalent of 75
From OBIAX: Math 1050 Grade vs Chem 1210 Pass Rate

Students Below 25 Math ACT:
Grade in Math 1050 vs 1210 Pass Rate

<table>
<thead>
<tr>
<th>Math 1050 Grade</th>
<th>1210 Pass Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>42%</td>
</tr>
<tr>
<td>D</td>
<td>44%</td>
</tr>
<tr>
<td>C</td>
<td>62%</td>
</tr>
<tr>
<td>B</td>
<td>78%</td>
</tr>
<tr>
<td>A</td>
<td>91%</td>
</tr>
</tbody>
</table>

Pre-Requisites:
- ‘C-’ or Better in Math 1050
- Others: ‘C-’ or Better in a Math Class Beyond 1050
- Score of 2 or Higher for AB or BC Calculus
Chem 1200: Prep for General Chemistry
- Semester-Long Course
- Taken BEFORE Chem 1210
- Basic Chemistry and Math Skills
- C- Set as a Course Pre-Requisite

Students Below 25 Math ACT:
Grade in 1200 vs 1210 Pass Rate

<table>
<thead>
<tr>
<th>1200 Grade</th>
<th>1210 Pass Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>38%</td>
</tr>
<tr>
<td>D</td>
<td>39%</td>
</tr>
<tr>
<td>C</td>
<td>74%</td>
</tr>
<tr>
<td>B</td>
<td>93%</td>
</tr>
<tr>
<td>A</td>
<td>97%</td>
</tr>
</tbody>
</table>
Pre-Requisite Summary

Accomplish One of the Following

Test Scores
- Math ACT: 25 or Greater
- Math SAT: 600 or Greater
- Math Accuplacer: 75 or Greater
- AB or BC Calculus: 2 or Greater

Math Courses: C- or Better
- Math 1050
- Another Math Course Beyond Math 1050

Chemistry Prep Course:
- Chem 1200: C- or Better
Improving Discussion Attendance

- Previous Years: Very Poor Discussion Attendance
- Recent Standardization of Discussion:
  - Multiple Choice and Numeric Response Questions
  - Fall 2013: Discussions Made as 5% Extra-Credit

Histogram of Discussion Percentages

<table>
<thead>
<tr>
<th></th>
<th>Fall 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>310</td>
</tr>
<tr>
<td>Mean</td>
<td>39.1%</td>
</tr>
<tr>
<td>Median</td>
<td>39.6%</td>
</tr>
<tr>
<td>St Dev</td>
<td>28.8%</td>
</tr>
</tbody>
</table>
Fall 2013: Determination of At-Risk Students in Chem 1220

- Determination of ‘At-Risk’
  - Previous Chem 1210 Performance
  - Use of Pre-Quiz the Beginning of the Semester

- At Risk If... (only one necessary)
  - <50% on Pre-Quiz
  - C+ or Below in Chem 1210

- Possibly At-Risk If... (both necessary)
  - B- in Chem 1210
  - 50-60% on Pre-Quiz

- Measure semester performance based on discussion attendance
Spring 2013: Discussion Attendance: Comparing At-Risk Students to those Not At-Risk

Findings:
Students we predict to be at-risk are very likely to never attend discussion when not required
Spring 2013: Pass Rate vs Discussion Attendance: Comparing At-Risk Students to those Not At-Risk

Findings:
Students Not At-Risk have only slight differences in pass rate based on discussion attendance
Students At-Risk dramatically increase in pass rate when regularly attending discussion
Historical Perspective Conclusion

- Students Not Likely to Attend Discussion, Even for Extra Credit
- At-Risk Students Benefit Most from Frequent Discussion Attendance
  - At-Risk Students Least Likely to Attend Discussion
- Result: Discussion Must Be Required
  - Fall 2014: Discussion Became 10% of Total Grade

End of Semester Course Enrollment
- Fall 2012: 1025
- Fall 2013: 999
- Fall 2014: 925
- Implementation of Pre-Requisites
- 7.5% Decrease in Enrollment 2013 to 2014
  - Likely the Result of Pre-Requisite Implementation
Discussion Results

Histogram of Discussion Percentages by Year

<table>
<thead>
<tr>
<th>Year</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>St Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2013</td>
<td>310</td>
<td>39.1</td>
<td>39.6</td>
<td>28.8</td>
</tr>
<tr>
<td>Fall 2014</td>
<td>299</td>
<td>75.1</td>
<td>84.1</td>
<td>25.6</td>
</tr>
</tbody>
</table>

One Class Used

- t-Test Results:
  - Statistically Different
  - $p < 0.001$

34.0% Increase in Discussion Percent
Histogram of Course Percentages

Histogram of Course Percent Scores by Year

<table>
<thead>
<tr>
<th>Year</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>St Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2013</td>
<td>639</td>
<td>69.9</td>
<td>72.2</td>
<td>18.4</td>
</tr>
<tr>
<td>Fall 2014</td>
<td>628</td>
<td>75.9</td>
<td>80.1</td>
<td>17.3</td>
</tr>
</tbody>
</table>

Two Classes of Data

\[ t\text{-Test Results:} \quad \text{Statistically Different} \]

\[ p < 0.001 \]

6.0% Increase in Course Percentage
Pass Rate

- Calculated Using the Standardized Rubric for Each Semester
- Based on a Course Total Percentage Greater than 69%
  - Fall 2013: 59.5%
  - Fall 2014: 72.3%
  - Change: 12.8% Increase in Pass Rate!
Where Is the Effect Occurring?

- Green: Passing
- Yellow: Not Passing, Within 1.5 SD of the Mean (based on Fall 2013): ‘Barely Failing’
- Red: Not Passing, Outside 1.5 SD of the Mean (based on Fall 2013): ‘Very Failing’

Conclusion: The ‘low hanging fruit’ of students barely failing are most responsive to positive course changes
**Histogram of Final Exam**

**Histogram ACS Final Normalized Scores by Year**

- Excluding those Not Taking the ACS Final:
  - Fall 2013: Percentile Median: 72, 74, 76
  - Fall 2014: Percentile Median: 79, 79, 81

### Two Class Results

- t-Test Results: Statistically Different
  - $p < 0.001$

<table>
<thead>
<tr>
<th></th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>639</td>
<td>628</td>
</tr>
<tr>
<td>Mean</td>
<td>46.6</td>
<td>68.0</td>
</tr>
<tr>
<td>Median</td>
<td>45</td>
<td>79</td>
</tr>
<tr>
<td>St Dev</td>
<td>32.8</td>
<td>30.6</td>
</tr>
</tbody>
</table>

**THE UNIVERSITY OF UTAH**
Implementation of Required Discussions

Conclusion

- Discussion Made Required
  - Result: Discussion Attendance Significantly Increased
- Pre-Requisites Implemented
  - Enrollment Somewhat Decreased
- Other Results
  - Pass Rate and Average Course Percent Significantly Increased
  - Standardized Score on ACS Exam Significantly Increased
Future Directions

- Creation of a Placement Exam for Chem 1210
  - Students placed in Chem 1210 or Chem 1200 depending on Score

- Current Ability:
  - 13 Question Quiz Created Measuring Student Problem Solving Ability

Spring 2015 1210 Final Percent vs Quiz Score

Statistically Significant Trend
\[ p < 0.001 \]
Future Directions

• Future Quiz Goals
  • Addition of Questions in the Following Categories
    • Logical Thinking
    • Chemistry Misconceptions
    • Math Ability
    • Chemistry Pre-Knowledge
  • Selection of the Best Questions

Prediction Ability Compared
• CCDT $R^2$: 0.17
• Our Current $R^2$: 0.20
• Goal $R^2$: 0.30 or Greater
• Students don’t know what they don’t know
  • Ability and metacognition linked\(^1\)

• Poor students in particular are overconfident\(^3\)

• Objective: Make students aware of their current ability

Does Ability to Predict Change Over Time?

• 15-Week Course with Multiple Tests
  • Students split into categories by final grade
  • Each Test: Prediction on Test Score
  • Students received test scores after each test

• Results
  • Good Students: Became More Accurate Over Time at Predicting Ability
  • Poor Students: Did not change in accuracy of predictions over time

Can Prediction Accuracy be Improved?

• Training Students Across the Semester
• End of Every Class Period: Exercises to Improve Monitoring Skills
  • Rated Confidence in Content Understanding
  • Describe the concept that was most difficult
• Practice questions about course content
  • Answered and Reported Confidence Judgments
    • How accurate are your answers
• Provided confidence judgments for each exam
  • First Test: Same as control group
  • Second Test: Improvement in judgment accuracy
  • Second to Final: One-Standard Deviation above Control in accuracy
• Better ability to measure understanding correlated with higher scores

Presentation Times vs Judgment of Learning

Underconfidence with Practice Effect

Cycle: 1) Studied a Topic  
2) Judgment of Learning on that Topic  
3) Tested on that Topic

Number of Presentations: How Many Times the Cycle was Repeated for the Topic

The Effect of Feedback

• Students Completed 11 One-Hour Tests
  • 200 General-Knowledge Questions with Two-Possible Answers
  • Selection of Answer
  • Judgment of Correctness: 50% Likely to 100% Likely as Correct
    • Most Participants were Overconfident Initially
  • Considerable Performance Feedback: Various Calibration Measures
  • Re-Tested
    • Subsequent Tests: Almost No Overconfidence
    • Calibration of Judgment with Only One Session of Training!

Pre-Test Loop Map

Homework Pre-Test → Study Plan → Homework Post-Test → Study Plan → Homework Score when Complete

Requirements Not Met

Requirements Met

Each Test
- Predict Ability
- Take the Test
- Postdict Ability by Topic
- IRT Analysis by Topic
- Feedback
- Study Plan
Future Directions Conclusion

• Students who do poorly are unaware of their level of ability
  • Poor students do not naturally improve in their monitoring of their ability over time
• When students regularly practice monitoring skills, student knowledge of their ability improves
  • Students’ ability to monitor their performance correlated with improved performance
• Repeated cycles of studying, assessment of ability, and testing: improves student ability and monitoring of ability
• Detailed feedback of students’ ability by topic rapidly improves student monitoring ability
Acknowledgements

• Henry White and Cynthia Burrows
  • Department Chairs
• Ronald and Eileen Ragsdale
• Nalini Nadkarni and Jordan Gerton
  • CSME Directors
Test Prediction/Post-Diction Details for Pre-, Post-, and Unit Tests

1) Score Prediction
   a. Sliding Scale: Marker Stops at every 5% between 0% and 100%
   b. Question: What Percent Do You Predict You Will Receive on the Test?
   c. Students Slide Marker to Predicted Test Score

2) Overall Test Prediction
   a. Question: How well do you feel you will do on the test compared to the rest of the class
      i. Likert Scale Options: Well Below Average, Below Average, Average, Above Average, Well Above Average

3) Topic Ability Prediction
   a. Topics: Problem Solving Ability or Conceptual Understanding by Chapter
      i. List all subtopics in each topic
   b. Question: How well do you feel you understand [topic] compared to the rest of the class
      i. Likert Scale Options: Well Below Average, Below Average, Average, Above Average, Well Above Average
<table>
<thead>
<tr>
<th>Q</th>
<th>Type of Problem</th>
<th>Intro Ch</th>
<th>Ch 1</th>
<th>Ch 2</th>
<th>Knowledge/Comprehension</th>
<th>Application</th>
<th>Unit Conversion</th>
<th>Calculations</th>
<th>Periodic Table</th>
<th>Energy and Light</th>
<th>Quantum Chemistry</th>
<th>Periodic Trends</th>
<th>Early Chemistry Laws</th>
<th>Measurement</th>
<th>Atomic Theory and Matter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Goal: predict what each student struggles with or understands based on patterns of what they get correct and incorrect.
Pre-tests

- Question pool reflective of topics on midterm exam

- Random non-repeating selection within topic for each pre-test
  - A student will see all questions once over four pre-tests

- Each student’s pre-test different
  - Advantage of IRT
IRT Analysis Overall

- Run IRT analysis on all questions using Bilog-MG
  - Overall question parameters (MMLE)
  - Overall student abilities (MLE)

**Question pool**
- thermo1
- thermo2
- thermo3
- thermo4
- acid/base1
- acid/base2
- acid/base3
- acid/base4
- equilibrium1
- equilibrium2
- equilibrium3
- equilibrium4

**Pre-test 1**
- thermo3
- acid/base4

**Student 1**
- thermo1
- equilibrium2
- thermo4

**All student responses**
- Student 1: Overall Ability
- Question Parameters

**Overall Ability**
IRT Topic Analysis

- Sort questions by topic
- IRT analysis of individual topics
  - Use only questions from topic
  - Student topic abilities
IRT Topic Analysis

Question pool
- thermo1
- thermo2
- thermo3
- thermo4
- acid/base1
- acid/base2
- acid/base3
- acid/base4
- equilibrium1
- equilibrium2
- equilibrium3
- equilibrium4

Student 1

thermo ability

- thermo3
- thermo1
- acid/base4
- acid/base1
- equilibrium2
- equilibrium4
IRT Topic Analysis

Question pool

thermo1
thermo2
thermo3
thermo4
acid/base1
acid/base2
acid/base3
acid/base4
equilibrium1
equilibrium2
equilibrium3
equilibrium4

Student 1
acid/base ability

thermo3  thermo1
acid/base4  acid/base1
equilibrium2  equilibrium4
IRT Topic Analysis

Question pool
- thermo1
- thermo2
- thermo3
- thermo4
- acid/base1
- acid/base2
- acid/base3
- acid/base4
- equilibrium1
- equilibrium2
- equilibrium3
- equilibrium4

Student 1
- equilibrium2
- equilibrium4

equilibrium ability
Individual Topic Abilities

- Topic abilities of twelve students on Fall 2014 midterm exam
- Individual students’ strengths and weaknesses
- Feedback most useful to students with high variant topic abilities

![Table of individual topic abilities]

Positive (blue) = high ability
Negative (red) = low ability
• Convert abilities to Likert scale
  • Well above average, above average, etc.

• Automated emails to individual students
  • Overall Likert ability
  • Likert ability for each topic
  • Likert ability for each question type
Feedback Report: Pre-, Post-, and Unit Tests

<table>
<thead>
<tr>
<th>Topics to Cover:</th>
<th>Prediction</th>
<th>Post-Diction</th>
<th>Actual Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Score</td>
<td>Score from Sliding Scale</td>
<td>Score from Sliding Scale</td>
<td>Percent on Test</td>
</tr>
<tr>
<td>Test Ability</td>
<td>Likert Scale</td>
<td>Likert Scale</td>
<td>IRT Likert Scale</td>
</tr>
<tr>
<td>Topic Ability (all topics listed)</td>
<td>Likert Scale</td>
<td>Likert Scale</td>
<td>IRT Likert Scale</td>
</tr>
</tbody>
</table>

- Sent to students through the program
- Report will include the student predicted, postdicted, and actual score or ability by area
Study Plan

1) Students check boxes within topic to create a study plan by topic
   1) Conceptual Ability by Chapter: Study Options
      1) Review In-Class Chapter Slides: Five Word Summary of Every Slide
      2) Read Chapter: Five Word Summary of Every Paragraph
      3) Concept Map of Chapter
      4) Outline of Chapter
      5) End-of-Chapter Conceptual Questions
   2) Problem-Solving Ability by Chapter
      1) Re-work in-class clicker questions by chapter
      2) Re-work discussion clicker questions by chapter
      3) Re-work homework questions by chapter
      4) End-of-chapter questions