FACILITATING CASE REUSE DURING PROBLEM SOLVING IN ALGEBRA-BASED PHYSICS

by

FRANCES ANN MATEYCIK

B.S., Rensselaer Polytechnic Institute, 2005

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department Of Physics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2009
Abstract

This research project investigates students’ development of problem solving schemata while using strategies that facilitate the process of using solved examples to assist with a new problem (case reuse). Focus group learning interviews were used to explore students’ perceptions and understanding of several problem solving strategies. Individual clinical interviews were conducted and quantitative examination data were collected to assess students’ conceptual understanding, knowledge organization, and problem solving performance on a variety of problem tasks.

The study began with a short one-time treatment of two independent, research-based strategies chosen to facilitate case reuse. Exploration of students’ perceptions and use of the strategies lead investigators to select one of the two strategies to be implemented over a full semester of focus group interviews. The strategy chosen was structure mapping.

Structure maps are defined as visual representations of quantities and their associations. They were created by experts to model the appropriate mental organization of knowledge elements for a given physical concept. Students were asked to use these maps as they were comfortable while problem solving. Data obtained from this phase of our study (Phase I) offered no evidence of improved problem solving schema. The 11 contact hour study was barely sufficient time for students to become comfortable using the maps.

A set of simpler strategies were selected for their more explicit facilitation of analogical reasoning, and were used together during two more semester long focus group treatments (phase II and phase III of this study). These strategies included the use of a step-by-step process aimed at reducing cognitive load associated with mathematical procedure, direct reflection of principles involved in a given set of problems, and the direct comparison of problem pairs designed to be void of surface similarities (similar objects or object orientations) and sharing physical principles (conservation of energy problems).
Overall, our results from the final two phases of this project indicate that these strategies are helpful in facilitating student ability to identify important information from given problems. The promising results from our study have significant implications for further research, curriculum material development, and instruction.
FACILITATING CASE REUSE DURING PROBLEM SOLVING IN ALGEBRA-BASED PHYSICS

by

FRANCES ANN MATEYCIK

B.S., Rensselaer Polytechnic Institute, 2005

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department Of Physics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2009

Approved by:

Major Professor
Dr. N. Sanjay Rebello
Abstract

This research project investigates students’ development of problem solving schemata while using strategies that facilitate the process of using solved examples to assist with a new problem (case reuse). Focus group learning interviews were used to explore students’ perceptions and understanding of several problem solving strategies. Individual clinical interviews were conducted and quantitative examination data were collected to assess students’ conceptual understanding, knowledge organization, and problem solving performance on a variety of problem tasks.

The study began with a short one-time treatment of two independent, research-based strategies chosen to facilitate case reuse. Exploration of students’ perceptions and use of the strategies lead investigators to select one of the two strategies to be implemented over a full semester of focus group interviews. The strategy chosen was structure mapping.

Structure maps are defined as visual representations of quantities and their associations. They were created by experts to model the appropriate mental organization of knowledge elements for a given physical concept. Students were asked to use these maps as they were comfortable while problem solving. Data obtained from this phase of our study (Phase I) offered no evidence of improved problem solving schema. The 11 contact hour study was barely sufficient time for students to become comfortable using the maps.

A set of simpler strategies were selected for their more explicit facilitation of analogical reasoning, and were used together during two more semester long focus group treatments (Phase II and Phase III of this study). These strategies included the use of a step-by-step process aimed at reducing cognitive load associated with mathematical procedure, direct reflection of principles involved in a given set of problems, and the direct comparison of problem pairs designed to be void of surface similarities (similar objects or object orientations) and sharing physical principles (conservation of energy problems).
Overall, our results from the final two phases of this project indicate that these strategies are helpful in facilitating student ability to identify important information from given problems. The promising results from our study have significant implications for further research, curriculum material development, and instruction.
Table of Contents

List of Figures .................................................................................................................................................. xv
List of Tables .................................................................................................................................................. xxii
Acknowledgements ........................................................................................................................................ xxv
Dedication ..................................................................................................................................................... xxvi

CHAPTER 1 - INTRODUCTION .................................................................................................................. 1
  1.1 Scope of Research ..................................................................................................................... 1
  1.2 Rationale ......................................................................................................................................... 2
  1.3 Research Questions ..................................................................................................................... 3
  1.4 Research Strategy Overview ..................................................................................................... 3
  1.5 Road Map of Dissertation ......................................................................................................... 5

CHAPTER 2 - LITERATURE REVIEW ........................................................................................................ 7
  2.1 Problem Solving Theory .............................................................................................................. 7
  2.1.1 Cognitive Processing ......................................................................................................... 7
  2.1.2 Expert-Novice Characteristics .......................................................................................... 8
    Problem Categorization ................................................................................................................... 9
    Problem categorization linked to schema acquisition and automation. ...................................... 10
  2.1.3 Transfer of Learning ........................................................................................................... 11
    Traditional Models .......................................................................................................................... 12
    Contemporary Models .................................................................................................................... 12
    Consolidating Contemporary and Traditional Perspectives of Transfer ..................................... 14
  2.1.4 Use of Analogy .................................................................................................................... 17
    Case-Based Reasoning ................................................................................................................... 17
    Lakoff's theory of metaphor ......................................................................................................... 17
    Tversky’s Contrast Model ............................................................................................................. 18
    Gentner’s rules for analogy: structure mapping ......................................................................... 19
    Podolofsky’s Model of Analogical Scaffolding ......................................................................... 22
    Case Reuse ..................................................................................................................................... 22
  2.2 Research based Problem Solving Strategies .............................................................................. 23
Focus Group Learning Interview 3 ................................................................. 70
Focus Group Learning Interview 4 ................................................................. 72
Focus Group Learning Interview 5 & 6 .......................................................... 76
Focus Group Learning Interview 7 & 8 .......................................................... 81
Focus Group Learning Interview 9 ................................................................. 86

4.3.2 Summary of FOGLI Results ................................................................... 88
How do students perform on the problems sets given in the focus group learning interviews? .......................................................................................................................... 89
How do students compare and contrast problems of varying similarity? .......... 89
How useful were the individual maps perceived while problem solving? .......... 89
How do structure maps affect collaborative problem solving? ....................... 90

4.4 Methodology – Individual Interviews ....................................................... 90
4.4.1 Mid-semester individual interview ....................................................... 92
Task 1 – Problem Posing .................................................................................. 92
Task 2 – text editing .......................................................................................... 93
Task 3 – Physics Jeopardy ................................................................................ 93
Task 4 – Creation of a concept map .................................................................. 93
4.4.2 End-semester individual interview ....................................................... 94
Task 1 – Problem Posing .................................................................................. 95
Task 2 – Physics Jeopardy ................................................................................ 95
Task 3 – Text Editing ....................................................................................... 96
Task 4 – Case reuse text editing ....................................................................... 96

4.5 Results - Individual Interview ................................................................ 97
4.5.1 Individual Interview 1 ........................................................................... 97
Task 1 – Problem Posing .................................................................................. 97
Task 2 – Text Editing ....................................................................................... 98
Task 3 – Physics Jeopardy ................................................................................ 99
Task 4 – Creation of a Concept Map ............................................................... 100
4.5.2 Individual Interview 2 ........................................................................... 101
Task 1 – Problem Posing .................................................................................. 101
Task 2 – Physics Jeopardy ................................................................................ 102
6.2 Screening Interviews........................................................................................................ 209
   6.2.1 Results from screening interviews.............................................................................. 209
6.3 Focus Group Learning Interviews.................................................................................. 212
   6.3.1 Methodology for Focus Group Learning Interviews................................................. 212
   6.3.2 Focus Group learning Interview – Week 1.............................................................. 216
       Methodology.................................................................................................................. 216
       Results for Week 1......................................................................................................... 217
   6.3.3 Focus group learning interview – Week 2............................................................... 219
       Methodology.................................................................................................................. 219
       Results for Week 2......................................................................................................... 220
   6.3.4 Focus group learning interview – Week 3.............................................................. 222
       Methodology.................................................................................................................. 222
       Results for Week 3......................................................................................................... 223
   6.3.5 Focus group learning interview – Week 4.............................................................. 227
       Methodology.................................................................................................................. 227
       Results for Week 4......................................................................................................... 228
   6.3.6 Focus group learning interview – Week 5.............................................................. 232
       Methodology.................................................................................................................. 232
       Results for Week 5......................................................................................................... 234
   6.3.7 Focus group learning interview – Week 6.............................................................. 238
       Methodology.................................................................................................................. 238
       Results for Week 6......................................................................................................... 239
   6.3.8 Summary of Results............................................................................................... 242
6.4 Individual Interviews..................................................................................................... 244
   6.4.1 Similarity Rating Task............................................................................................... 246
   6.4.2 Results – Similarity Rating Task .............................................................................. 247
   6.4.3 Usability Task............................................................................................................ 250
   6.4.4 Results – Usability Task........................................................................................... 251
6.5 Non-Traditional Problems on Examinations................................................................ 257
   6.5.1 Text Editing .............................................................................................................. 258
   6.5.2 Physics Jeopardy...................................................................................................... 259
A-4 Interview 2 protocols ................................................................. 328
A-5 Institutional Review Board consent forms ................................ 332
APPENDIX B - Phase I ................................................................. 335
   B-1 Structure Maps ................................................................. 335
   B-2 Focus Group Learning interview protocols ......................... 349
   B-3 Individual Interviews ....................................................... 394
   B-4 Institutional Review Board consent forms .......................... 402
APPENDIX C - Phase II ................................................................. 407
   C-1 Screening Interview Protocol .......................................... 407
   C-2 Focus Group Learning Interviews ..................................... 410
   C-3 Individual Interview 1 ..................................................... 564
   C-4 Individual Interview 2 ..................................................... 585
   C-5 Problem files for Individual Interview 1 & 2 .................... 607
   C-6 Extra-credit examinations .............................................. 620
   C-7 Institutional Review Board consent forms ....................... 625
APPENDIX D - Phase III ............................................................. 633
   D-1 Screening Interview protocols ........................................ 633
   D-2 Focus Group Learning Interviews ..................................... 636
   D-3 Individual Interview protocol .......................................... 766
   D-4 Problem files for Individual Interview ............................. 782
   D-5 Extra Credit for Examinations ........................................ 789
   D-6 Institutional Review Board consent forms ....................... 794
APPENDIX E - Statistics for Phase II and III ................................. 800
List of Figures

Figure 2.1 The mechanism of transfer proposed by Rebello et. al., 2005 ....................... 15
Figure 2.2 Tversky’s contrast model. .............................................................................. 19
Figure 2.3 Structure map of Rutherford analogy ............................................................. 21
Figure 2.4 Work and energy structure map. ................................................................. 27
Figure 2.5 Sample problem posing sheet ..................................................................... 31
Figure 3.1 Questioning strategy example. ..................................................................... 36
Figure 3.2 Structure map and questions for work-energy problem pairs. ...................... 40
Figure 3.3 Work-Energy examination problem and solution. ......................................... 46
Figure 3.4 Work-Energy problem from interview 1. ...................................................... 50
Figure 4.1 Research design timeline. ............................................................................ 60
Figure 4.2 Top left: Problem 1 from week 6. Top right: Problem 2 from week 6. Bottom left: Problem 3 from week 6. Bottom right: Problem 4 from week 6. Problems above are variations of problems asked in chapter 6 of Physics: Principles with Applications, Giancoli, 6th Edition ................................................................. 62
Figure 4.3 Visual description of the variance of problem types given during the focus group learning interviews. ...................................................................................... 63
Figure 4.4 Left: A structure map used during week 4. Right: A structure map used during week 6. Problems selected for students to work on while using these maps were variations of problems asked in chapters 4 and 6 of Physics: Principles with Applications, Giancoli, 6th Edition, respectively. ......................................................... 64
Figure 4.5 FOGLI 1 problem set ................................................................................. 66
Figure 4.6 An example of the FOGLI 2 problems sets ................................................... 69
Figure 4.7 Problem set and force structure map used for FOGLI 3. ................................. 71
Figure 4.8 FOGLI 4 structure map and problem set ..................................................... 73
Figure 4.9 Work and Energy structure map and problem set for FOGLI 5 .................... 77
Figure 4.10 FOGLI 6 problem set ................................................................................ 77
Figure 4.11 Example of students' map markings (darker color represents the difference in quantities between two maps). ............................................................................. 79
Figure 4.12 Vibration and wave structure map and problem set for FOGLI 7 ............. 82
Figure 5.21 Diagram of possible problem pairings with variances of similarities and
differences.......................................................................................................................... 157
Figure 5.22 Example of a problem sequence for the identification task. .................. 158
Figure 5.23 Type A problem pair: Facial similarity [FS] (both roller coasters) and
principle similarity [PS] (both are conservative systems). ......................................... 165
Figure 5.24 Type B problem pair: Facial difference [FD] (roller coaster vs. gun) and
principle similarity [PS] (both are conservative systems). ........................................... 165
Figure 5.25 Type C problem pair: Facial similarity [FD] (both roller coaster) and
principle difference [PD] (conservative vs. non-conservative). ................................. 165
Figure 5.26 Type D problem pair: Facial difference [FD] (roller coaster vs. gun) and
principle difference [PD] (both are conservative systems). ........................................... 166
Figure 5.27 Students' similarity ratings of problem pairs of type A, B, C, and D for
interview 1 and interview 2............................................................................................. 167
Figure 5.28 Challenging problem for individual interview 1. ...................................... 172
Figure 5.29 Challenging problem given for individual interview 2. ............................ 172
Figure 5.30 Timeline showing when exams were administered during the semester
relative to the FOGLI's..................................................................................................... 188
Figure 5.31 Example of a multiple choice adaptation of a text editing task................. 189
Figure 5.32 Example of a multiple choice adaptation of physics Jeopardy task......... 190
Figure 5.33 Example of a multiple choice adaptation of a problem posing task........ 191
Figure 5.34 Split plot of the baseline and treatment groups average scores on traditional
problems for each examination..................................................................................... 200
Figure 5.35 Split plot of the baseline and treatment groups average scores on Physics
Jeopardy problems for each examination. ................................................................. 201
Figure 5.36 Split plot of the baseline and treatment groups average scores on text editing
problems for each examination..................................................................................... 202
Figure 5.37 Split plot of the baseline and treatment groups average scores on problem
posing problems for each examination........................................................................ 203
Figure 6.1 Research timeline for the replication of the focus group learning interviews.
........................................................................................................................................ 208
Figure 6.2 Example problem C and problems A and B from week 4......................... 213
Figure 6.3 Example of step procedure given during the focus group learning interviews.  
................................................................................................................................. 214

Figure 6.4 Example of procedural scaffolding. .............................................................. 215

Figure 6.5 Example of table given for similarities and differences and rating scale for the usability rating. ................................................................. 216

Figure 6.6 Problems used during FOGLI 1. ................................................................. 217

Figure 6.7 Problems used during FOGLI 2. ................................................................. 219

Figure 6.8 Problems used during FOGLI 3. ................................................................. 223

Figure 6.9 Example of problem posed during week 3. ................................................... 227

Figure 6.10 Problems used during FOGLI 4. ................................................................. 227

Figure 6.11 Example of a problem pair given in week 4. ................................................. 232

Figure 6.12 Problems used during FOGLI 5. ................................................................. 233

Figure 6.13 Example of a problem posed given in week 5. ............................................ 237

Figure 6.14 Problems used during FOGLI 6. ................................................................. 238

Figure 6.15 Example of a problem pair given in week 6. ................................................ 242

Figure 6.16 Diagram representing two scenario sets. ..................................................... 245

Figure 6.17 Example of possible problem pairs with variances of similarities and difference. ............................................................................................................. 246

Figure 6.18 Similarity ratings for the individual interview in phase III. ............. 248

Figure 6.19 Similarity ratings for phase II, interview 2 and phase III ..................... 249

Figure 6.20 Faculty and end interview ratings from Phase II and Phase III ........... 250

Figure 6.21 Challenging problem used in the usability task ........................................ 251

Figure 6.22 Timeline for examinations and FOGLI's .................................................... 258

Figure 6.23 Example of a multiple choice adaptation of a text editing task ............ 258

Figure 6.24 Example of a multiple choice adaptation of physics Jeopardy task ........ 259

Figure 6.25 Example of a multiple choice adaptation of a problem posing task ........ 260

Figure 6.26 Text editing task on examination 3 ............................................................. 264

Figure 6.27 Split plot of the baseline and treatment groups average scores on traditional problems for each examination ................................................. 270

Figure 6.28 Split plot of the baseline and treatment groups average scores on Physics Jeopardy problems for each examination ................................................. 271
Figure 6.29 Split plot of the baseline and treatment groups average scores on text editing problems for each examination. ................................................................. 271

Figure 6.30 Split plot of the baseline and treatment groups average scores on problem posing problems for each examination. .................................................. 272

Figure A.1 Work-kinetic energy structure mapping strategy. ........................................ 300

Figure A.2 Potential energy structure mapping strategy. ............................................ 301

Figure A.3 Treatment Assessment problems. ............................................................... 305

Figure A.4 Control Group Problems. .......................................................................... 306

Figure A.5 Statistical results stemming from the extra credit assignments. ............... 307

Figure A.6 Work and energy examination problem. ...................................................... 319

Figure A.7 Work-Energy problem for interview 1. ...................................................... 327

Figure A.8 Work and energy problem (same as used for interview 1). ....................... 331

Figure A.9 IRB Form for spring 2007 individual interviews. ....................................... 333

Figure B.1 The structure map training guide. ................................................................ 335

Figure B.2 Force maps. ............................................................................................... 344

Figure B.3 Work and energy map. ............................................................................... 346

Figure B.4 Vibrations and waves maps. ...................................................................... 347

Figure B.5 Week 1 protocol. ....................................................................................... 349

Figure B.6 Week 2 protocol including solutions. ......................................................... 354

Figure B.7 Week 3 protocol including solutions. ......................................................... 360

Figure B.8 Week 4 protocol including solutions. ......................................................... 365

Figure B.9 Week 5 protocol with solutions. ................................................................. 371

Figure B.10 Week 6 protocol with solutions. ............................................................... 375

Figure B.11 Week 7 protocol with solutions. ............................................................... 379

Figure B.12 Week 8 protocol with solutions. ............................................................... 383

Figure B.13 Week 9 protocol with solutions. ............................................................... 387

Figure B.14 Individual interview 1 protocol (mid-semester interview). .................... 394

Figure B.15 Individual interview protocol 2 (end-semester interview). ..................... 398

Figure B.16 IRB Consent forms for focus group learning interview participants........ 403

Figure B.17 IRB Consent forms for individual interview participants.......................... 405

Figure B.18 Volunteer slip given to students during lecture. ....................................... 406
Figure C.1 Screening interview protocol. ................................................................. 407
Figure C.2 Week 1 Summary: 02-18-08. ................................................................. 410
Figure C.3 Week 2 summary: 02-25-08. ................................................................. 425
Figure C.4 Week 3 summary: 03-03-08. ................................................................. 437
Figure C.5 Week 4 summary: 03-10-08. ................................................................. 454
Figure C.6 Week 5 summary: 04-07-08. ................................................................. 475
Figure C.7 Week 6 summary: 04-14-08. ................................................................. 492
Figure C.8 Week 7 summary: 04-21-08. ................................................................. 515
Figure C.9 Week 8 summary: 04-28-08. ................................................................. 541
Figure C.10 Individual interview 1: Identification Task - Set 1. ......................... 564
Figure C.11 Individual interview 1: similarity ratings task - Set 1. ....................... 569
Figure C.12 Individual interview 1: Identification task - Set 2. ......................... 574
Figure C.13 Individual interview 1: Similarity ratings task - Set 2. ....................... 579
Figure C.14 Individual interview 1: Usability task problem. ......................... 583
Figure C.15 Similarity ratings survey data collected from interview 1 worksheets. 584
Figure C.16 Individual interview 2: Identification task - Set 1. ......................... 585
Figure C.17 Individual interview 2: Similarity ratings task - Set 1. ....................... 591
Figure C.18 Individual interview 2: Identification task - Set 2. ......................... 595
Figure C.19 Individual Interview 2: Similarity ratings task - Set 2 ....................... 601
Figure C.20 Individual Interview 2: Usability task problem. ......................... 605
Figure C.21 Data collected from interview 2 worksheets. ................................. 606
Figure C.22 Problems used for individual interview 1. ................................. 607
Figure C.23 Problems used for individual interview 2. ................................. 613
Figure C.24 Problem printing sequence for interview 1. ................................. 619
Figure C.25 Problem printing sequence for interview 2. ................................. 619
Figure C.26 Extra credit for examination 1. ............................................................. 620
Figure C.27 Extra credit for examination 2. ............................................................. 621
Figure C.28 Extra credit for examination 3. ............................................................. 622
Figure C.29 Extra credit for examination 4. ............................................................. 623
Figure C.30 Extra credit for examination 5. ............................................................. 624
Figure C.31 IRB Consent form for screening interview participants. .................. 626

xx
Figure C.32 IRB Consent form for focus group learning interview participants .......... 628
Figure C.33 IRB Concent form for individual interview participants ........................... 630
Figure C.34 Volunteer slip collected during lecture...................................................... 632
Figure D.1 Screening interview protocol ................................................................. 633
Figure D.2 Week 1 Summary: 09-24-08. ................................................................. 637
Figure D.3 Week 2 summary: 10-02-08. ............................................................... 652
Figure D.4 Week 4 summary: 10-22-08. ............................................................... 694
Figure D.5 Week 5 summary: 11-05-08. ............................................................... 721
Figure D.6 Week 6 summary: 11-12-08. ............................................................... 744
Figure D.7 Individual interview: Similarity ratings task - Set 1............................... 766
Figure D.8 Individual interview: Usability task - Set 1............................................ 770
Figure D.9 Individual interview: Similarity ratings task - Set 2............................... 773
Figure D.10 Individual interview: Usability task - Set 2........................................... 777
Figure D.11 Individual interview: Usability task problem ....................................... 780
Figure D.12 Similarity ratings data collected from interview worksheets ............. 781
Figure D.13 Problems used for the individual interview ......................................... 782
Figure D.14 Problem printing sequence for the interview ..................................... 788
Figure D.15 Extra credit for examination 1 ............................................................. 789
Figure D.16 Extra credit for examination 2 ............................................................. 790
Figure D.17 Extra credit for examination 3 ............................................................. 791
Figure D.18 Extra credit for examination 4 ............................................................. 792
Figure D.19 Extra credit for examination 5 ............................................................. 793
Figure D.20 IRB Consent form for screening interview participants ..................... 795
Figure D.21 IRB Consent form for focus group and individual interview participants 797
Figure D.22 Volunteer slip collected during lecture .............................................. 799
List of Tables

Table 1.1 Case reuse project research timeline ................................................................. 5
Table 2.1 Example of a metaphor using resistance ......................................................... 18
Table 3.1 The questioning strategy group Pearson correlation values ............................ 43
Table 3.2 The structure mapping group Pearson correlation values ............................... 43
Table 3.3 The mean percentage correct on questions given during the structure mapping 
    and questioning strategy treatments ........................................................................ 43
Table 3.4 Grading rubric for all problems given as part of the extra-credit assignment . 44
Table 3.5 Student performance on problems assigned with extra-credit ....................... 45
Table 3.6 ANOVA problem performance significance between groups ..................... 45
Table 3.7 ANOVA calculation of final score significance between the three groups ...... 45
Table 3.8 Rubric to assess performance on the 'far transfer' examination problem ....... 47
Table 3.9 Rubric scores for treatment and control groups on the examination far transfer 
    problem .................................................................................................................... 49
Table 4.1 Demographics for phase 1 participants ......................................................... 61
Table 5.1 Demographics for Phase II participants ......................................................... 113
Table 5.2 Similarity and usefulness ratings for week 4 .................................................. 131
Table 5.3 Similarity and usefulness ratings for week 5 .................................................. 136
Table 5.4 Similarity and usefulness ratings for week 6 .................................................. 140
Table 5.5 Similarity and usability ratings for week 7 ..................................................... 145
Table 5.6 Similarity and usability ratings for the final week 8 ...................................... 149
Table 5.7 Problem pairs for the similarity rating task .................................................. 164
Table 5.8 Usability problems selected by students for set 1, interview 1 as most and least 
    useful ..................................................................................................................... 174
Table 5.9 Usability problems selected by students for set 2, interview 1 as most and least 
    useful ..................................................................................................................... 177
Table 5.10 Most and least useful problem selections categorized by surface feature in 
    individual interview 1 ........................................................................................... 179
Table 5.11 Usability problems selected by students for set 1, interview 2 as most and 
    least useful ............................................................................................................ 180
Table 5.12 Usability problems selected by students for set 2, interview 2 as most and least useful. ................................................................................................................................. 183
Table 5.13 Most and least useful problem selections categorized by surface feature in individual interview 2. ........................................................................................................... 187
Table 5.14 Comparison of scores of cohort and rest of class on traditional exam problems.............................................................................................................................. 193
Table 5.15 Comparison of scores of cohort and rest of class on text editing tasks. ...... 194
Table 5.16 Comparison of scores of cohort and rest of class on physics Jeopardy tasks. .............................................................................................................................. 195
Table 5.17 Comparison of scores of cohort and rest of class on problem posing tasks. 196
Table 5.18 Significance between the baseline and treatment groups on traditional and non-traditional problem scores averaged over all five examinations. ................. 198
Table 5.19 Significance of difference on average traditional and non-traditional problem scores between examinations.................................................................................. 198
Table 5.20 Significance in difference between the treatment and baseline groups on traditional and non-traditional problem scores between the five examinations..... 199
Table 6.1 Demographics for the 12 participants in this study. ....................................... 211
Table 6.2 Similarity and Usability Ratings for week 2.................................................... 222
Table 6.3 Similarity and usability ratings for week 3...................................................... 225
Table 6.4 Similarity and usability ratings for week 4...................................................... 231
Table 6.5 Similarity and usability ratings for week 5...................................................... 235
Table 6.6 Similarity and usability ratings for week 6...................................................... 240
Table 6.7 Problem pairs for the similarity rating task. ............................................... 247
Table 6.8 Usability problems selected by students for set 1 as most and least useful.... 252
Table 6.9 Student selections of problems for the individual interview. ....................... 254
Table 6.10 Usability problems selected by students for set 2 as most and least useful.. 255
Table 6.11 Comparison of scores of cohort and rest of class on traditional exam problems.............................................................................................................................. 262
Table 6.12 Comparison of scores of cohort and rest of class on text editing tasks. ...... 263
Table 6.13 Comparison of scores of cohort and rest of class on physics Jeopardy tasks. .............................................................................................................................. 265
Table 6.14 Comparison of scores of cohort and rest of class on problem posing tasks. 266
Table 6.15 Significance between the baseline and treatment groups on traditional and non-traditional problem scores averaged over all five examinations. 268
Table 6.16 Significance of difference on average traditional and non-traditional problem scores between examinations. 268
Table 6.17 Significance in difference between the treatment and baseline groups on traditional and non-traditional problem scores between the five examinations. 269
Table A.1 Treatments and pilot representation summary. 289
Table A.2 Work-kinetic energy questioning strategy. 291
Table A.3 Potential energy questioning strategy. 294
Table A.4 Conservation of energy questioning strategy. 297
Table A.5 Conservation of Energy Structure mapping strategy. 303
Table A.6 Problem rubric and codes. 320
Table A.7 Inter-rater reliability scores. 321
Table A.8 Examination results for exam 2. 323
Table A.9 Interview 1 Protocol. 326
Table A.10 Interview 2 protocol for structure mapping and questioning strategy. 328
Table A.11 Interview 2 protocol for control group. 329
Table E.1 Problem-by-problem ANOVA statistics for Phase II and III. 801
Table E.2 Exam-by-Exam and Exam-by-Treatment interaction ANOVA Statistics for Phase II and III. 803
Acknowledgements

This work has been supported in part by the National Science Foundation grant DUE-06185459.

I wish to acknowledge the support and guidance provided by my graduate advisor, Sanjay Rebello. This work would not be possible without his counseling and expertise.

I would also like to thank the following individuals for their assistance in my research efforts: David Jonassen, Zdeslav Hrepic, and Carlos Wexler for their essential contributions to the research design; Leigh Murray for her statistical expertise; Brett DePaola, Noel Stanton, and Gary Wysin for their willingness to permit us access to their students; my committee members: Amit Chakrabarti, Andrew Bennett, Lawrence Scharmann, and Oliver Weaver for their unwavering support.

My special thanks to Dean Zollman, Sanjay Rebello, and the Kansas State University Physics Education Research Group for their continual support and contributions to this work.
Dedication

To my father, John G. Mateycik, Jr.
CHAPTER 1 - INTRODUCTION

The development of an effective and efficient problem solving approach is an important goal in several introductory physics courses. Over the years, the physics education research community has devoted significant effort to research in problem solving (Hsu, Brewe et al., 2004). Previous studies, many described in Hsu et al.’s resource letter, investigate problem solving issues such as creation and use of mental representations, student cognitive load, and the effectiveness of problem solving strategies and heuristics on student performance. The same resource letter also demonstrates that many of these works are closely tied to a variety of disciplines including but not limited to cognitive science, education, chemistry education, mathematics education and biology education. Studies cited in the resource letter also include work that ties assessment of students’ conceptual understanding to ‘non-traditional’ problem solving tasks. This problem solving study, well-informed by these previous studies, focused on facilitating a common student problem solving approach known as case reuse (Jonassen, 2006). Case reuse may be defined as the process of using a previous example or solved problem to assist with the resolution of a new, but analogous problem.

In this chapter, I will present the foundations of this study, motivation for conducting this research and the overarching research questions which prompted the pilot and three phases of this project. I will also discuss our research strategy and conclude with an outline of the dissertation chapters.

1.1 Scope of Research

An important skill in problem solving is recognizing that a given problem can be solved using the same or similar approach as a problem with which one is already familiar. (Kolodner, 1997; Quilici and Mayer, 2002; Jonassen, 2003) For example, one might use a kinematics problem in their textbook to help solve another kinematics problem in their homework. The problems may not be identical, but many of the procedural and conceptual elements remain the same. Unfortunately, while students attempt to utilize their previously encountered examples, they may have difficulty with
selection of appropriate cases. Catrambone & Holyoak (Catrambone and Holyaok, 1989) and Reed (Reed, 1987) suggest that learners fail to recall examples appropriately because their retrieval is based upon similarity of objects between examples, not their structural features. Catrambone & Holyoak suggest that generalization improves when problems emphasize structural features shared with a similar example. Jonassen and Chi (Chi, Feltovich et al., 1981; Jonassen, 2003) further note that a scaffolding strategy or strategies may be necessary for strengthening students’ reflection of the conditions associated with physical concepts and principles. Thus, in our own project, we looked to identify strategies that sufficiently emphasize structural features and guide students’ reflection while using worked examples. This study contributes to the research on novice strategies and provides a framework for helping students learn problem solving more effectively using prior cases.

It is important to note that this project’s focus does not revolve around the time constraints or content covered over a given semester of introductory physics. There are three assumptions that carry with the future use of the intervention created and assessed in this project. Assumption 1: the course must allot some time span for students to discuss physics problems. Assumption 2: the educator must be willing to actively motivate collaborative student discussion and reflection during class hours. Assumption 3: one of the primary focuses of the course is to provide students with better problem solving skills.

1.2 Rationale

Ideally, the nature of human interest in scientific discovery would be enough to warrant any research question. But particularly for this project, why do we choose to spend our time (and money) looking for a way to facilitate a common, unreliable and inefficient novice problem solving approach?

The answer is fairly simple. Individuals, including experts of science and mathematics, commonly extrapolate information collected and stored from previous events to determine how it might be comparable to a new circumstance. Experts refine their approach to case reuse over years of experience. This refinement includes a more sophisticated organization of knowledge elements and their associations (Jonassen,
2006). Our goal is to assist novices with refining this approach early in their studies, easing both cognitive load and the perceived difficulty associated with physics problems. With the continual decline of student’s choosing physical science as either a major or even a science elective, research focus has turned toward students’ attitudes toward science and how the current decline might be reversed. (Osborne, Simon et al., 2004) This study will present evidence of how implementing strategies that accommodate students’ pre-existing problem solving methods positively affect student performance and their overall attitude towards physics problem solving.

1.3 Research Questions

This research project aims to address the following central research question:

■ What scaffolding facilitates case-reuse in problem solving with students in an introductory college algebra-based physics course?

The following sub questions arise from the question above:

■ What scaffolding -- cues, hints, activities and other external inputs -- cause students to reorganize their knowledge while problem solving?

■ To what extent can they utilize this scaffolding to reorganize their knowledge while solving problems?

■ What are the ways in which the expert-like strategy of asking productive questions about a problem can be assimilated in students’ problem solving repertoire?

■ To what extent is the strategy of facilitating students to ask expert-like questions an effective way to help students solve problems in physics?

■ To what extent do students’ attitudes about problem solving change after experiencing the problem solving strategies promoted in this project?

1.4 Research Strategy Overview

We primarily investigated strategies for facilitating case reuse in physics problem solving through phenomenographic analysis (Marton, 1986) of focus group learning interviews. These group interviews were conducted weekly with algebra-based students. During the focus group learning interviews, we looked to observe, describe, and
understand how students’ experienced the facilitating strategies. Since these strategies were dependent on explicit reflection, group collaborations were an effective way to establish active participation and discussion between learners. They were also a feasible activity for possible future classroom implementation. Assessment of students’ perceptions and performance was done using semi-structured individual clinical interviews, quantitative analysis of in-class examinations, and quantitative analysis of non-traditional problem tasks given as extra-credit during the in-class examinations.

For our initial pilot study, algebra-based physics students were given a one time, online treatment as an extra-credit activity. We investigated students’ performance by comparing our two treatment groups to a control group. The online treatment groups were assessed using quantitative data collected from the treatment questions and three extra-credit problems from the textbook. The control group was assessed using six extra-credit problems from the textbook. Three of those six were the same as those given to the treatment groups. Finally, all students participating in the treatment groups and control group were assessed on their performance on an examination word problem using an in-house, conceptually oriented grading rubric.

A thorough survey of problem solving literature was conducted to understand models of analogical reasoning and novice problem solving approaches prior to the selection of all strategies used in this project. It was concluded from the pilot and phase I, (described in chapters three and four, respectively), that certain strategies are difficult to implement in a one-time, or one semester term. Phase I of this study emphasized the importance of students’ perceptions of a given strategy, and how students’ affect their willingness to carry out such strategies on their own. Phase II and III, (described in chapters five and six, respectively), incorporate strategies which are less demanding for instructors or moderators and are more feasibly implemented in one semester. These strategies continue to facilitate students to reflect on the information contained within problem statements, but they also more explicitly direct attention to deep-structure (physics principles) similarities between cases.
1.5 Road Map of Dissertation

This dissertation consists of seven chapters. The following table summarizes the timeline for each phase of this project and how the project is outlined in this dissertation.

Table 1.1 Case reuse project research timeline.

<table>
<thead>
<tr>
<th>Chapters</th>
<th>Chapter Focus</th>
<th>Semester research was conducted:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Project outline</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>Literature Review</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>Pilot</td>
<td>Spring 2007</td>
</tr>
<tr>
<td>4</td>
<td>Phase 1</td>
<td>Fall 2007</td>
</tr>
<tr>
<td>5</td>
<td>Phase 2</td>
<td>Spring 2008</td>
</tr>
<tr>
<td>6</td>
<td>Phase 3</td>
<td>Fall 2008</td>
</tr>
<tr>
<td>7</td>
<td>Conclusions</td>
<td>-</td>
</tr>
</tbody>
</table>

This first chapter discusses the general scope of this project and the underlying motivation for conducting the research. Chapter 1 also contains the research questions, and an overview of the research strategy.

Chapter 2 provides a review of relevant research literature including theories related to cognitive processing, expert versus novice problem solving, dynamic transfer of learning, and the use of analogies. Chapter 2 also includes a review of problem solving strategies which aim to direct attention to the problem statement features important to physical phenomenon. Chapter 2 concludes with a review of assessments of conceptual understanding.

Chapter 3 discusses the research methodology, analysis, and results pertaining to the pilot study conducted with students from an algebra-based physics course. The chapter concludes with a discussion of the pilot’s resulting impact on the next phase of our research study.

Chapter 4 discusses the research methodology, analysis, and results pertaining to the first phase of our study conducted over one semester with algebra-based physics students. This is the first of three phases that use focus group learning interviews to gather data. Chapter 4 concludes with a discussion of the crucial transition from the first phase to the second phase of this research project.
Chapter 5 illustrates the research methodology, analysis, and results pertaining to the second phase of this research study. This phase was also conducted over a semester long treatment with algebra-based physics students. Chapter 6 presents a replication of the second phase using the finalized protocols created mid-way through the second phase. Chapter 5 concludes with a discussion of limitations and future study. Chapter 6 concludes with a comparison between the two semesters, and the promising possibilities implied by the congruent results.

Finally, chapter 7 summarizes the key findings from the different phases of this research project by answering our research questions. This dissertation concludes with a discussion of the implications of this research for further study including the possibility for full course integration.
CHAPTER 2 - LITERATURE REVIEW

In this chapter, I review the research literature related to problem solving. For our project, we looked to extend the understanding of how students use analogies while problem solving, and to what extent certain pre-existing problem solving strategies may facilitate analogical reasoning. We informed ourselves of the theoretical work with regards to mental problem solving processing, and how students transfer elements of their pre-existing knowledge structures to other contexts. These theoretical representations of organized knowledge and the dynamic model of transfer of learning are presented first. The literature review continues with theories of students’ use of analogy, including case reuse, the work in which this project is primarily based. Finally, problem solving strategies and assessments which show promise of lending measurable attention to students’ problem solving schemata are explained. The chapter concludes with a discussion of relevant research frameworks and their extension to our own project.

2.1 Problem Solving Theory

2.1.1 Cognitive Processing

Problem solving can involve several mental tasks: recall of relevant information from memory, creating representations from information given in a problem (force-diagrams or hierarchical structures), checking for consistency (checking one’s own work for contradictions in mathematical or physical logic), and others. (Hsu, Brewe et al., 2004) For our purpose, as we align to augment student’s problem solving skills, we need to better understand the process involved during problem solving. Anderson et. al. (Anderson, Greeno et al., 1981) discuss the processes of acquisition, compilation, and optimization of cognitive skills while solving geometry proof problems. According to Anderson and colleagues, students text learn while they encode from the text of the problem and process the usefulness of the encoded information for problem solving. The process of subsumption occurs at this time and is the means by which students encode information into existing knowledge structures. Students then must transform encoded information, which is declarative, to a more effective procedural form before beginning the planning stage. For example, Anderson et. al. determined that attempts at geometric
proof generation can be divided into two major stages: student attempts to find a plan for the proof (planning mode) and the student translation of that plan into an actual proof (execution mode). Students were often observed alternating back and forth between the two modes when discrepancies came to pass during the translation of the intended plan into a coherent proof. Anderson’s conclusion from this detailed study of students’ process of generating proofs is that practice does in fact help students build better knowledge structures.

Reif and Heller (Reif and Heller, 1982) divide the process of a problem solving into three phases: the description phase, the search-for-a-solution phase, and the assessing-the-solution phase. The description phase is when a solver breaks the problem statement into a clear description of the problem and the information to be found. The search for a solution phase may be facilitated by decomposition or constraint satisfaction. While searching for a solution one might produce a list of constraints with respect to information obtained in the problem statement and create a solution that satisfies the constraints (constraint satisfaction). One might also break a problem into progressively ‘easier to solve’ parts (decomposition). The final phase of assessing a solution is essentially ensuring that the solution is complete, consistent, and optimal. All of these phases are fairly general and provide a basic sequence to problem solving. In this same paper, Reif and Heller also discuss the need for creation of new representations of a problem within every new phase in the problem solving procedure. It is suggested from Reif’s earlier work with Jill Larkin (Larkin and Reif, 1979) that eliciting formal planning of a solution prior to working out the mathematical detail may better mimic expert behavior. They asked an expert and a novice to solve five problems while thinking aloud. They determined that the expert constructed a "low-detailed qualitative physical description" after a sketch. This was done as a self-consistency check, insuring that their determination of a proper approach to a solution is not ill-considered. The novice did not construct a similar description. Larkin and Reif, concluded that experts “rapidly re-describe problems presented to them, often use qualitative arguments to plan solutions before elaborating them in greater mathematical detail, and make many decisions by first exploring their consequences.”

2.1.2 Expert-Novice Characteristics
Our goal in this project is to enhance student organization of their knowledge using effective reasoning strategies. To accomplish this goal, we must try to understand how experts differ from novices with regards to their problem solving approaches. We consider some fundamental studies regarding expert and novice problem solving below.

**Problem Categorization**

One of the most cited and well-known works examining how experts structure their knowledge was conducted by Chi, Feltovich and Glaser (Chi, Feltovich et al., 1981). According to Chi and her colleagues, the expert's knowledge base is arranged around a problem schema, or mental organization containing information necessary to solve a specific category of problems. To uncover information regarding expert and novice differences in these categories, Chi conducted several studies. The first study required eight Physics Ph.D. candidates and eight undergraduates to sort 24 problems based on similarity in solutions. Chi et al. observed that novices categorized problems based on the surface features of the problem. Surface features were defined as particular objects or physical configurations referred to in a problem. The experts did not use these surface features for their categorization. Experts categorized based on the major physics principles used in the solution to the problem. A second study, using a similar sampling, was conducted using problems which had similar surface features, but different physical principles necessary for a solution. In this study they confirmed that novices use surface features.

By the third study, Chi focused her effort on examining the difference in problem schemata held by experts and novices. Two experts and two novices were given 20 category labels composed of physical principles and surface features. The subjects then had three minutes to describe all they knew about problems involving each category. The researchers concluded that experts associated principles with procedural knowledge about their applicability. Chi et al. inferred that the experts’ problem schemata’s contained an ordered hierarchy of conditions associated with physical principles. The novices focused on surface details and expressed details about finding explicit unknowns in a given problem. Chi determined that novices’ schemata were lacking information with respect to the application of principles. For the fourth study, two experts and two novices were
invited to think aloud while solving 20 problems. This fourth study, showed results similar to Larkin and Reif, where experts expressed the underlying principles involved in a problem, while novices jumped directly into solving the problems mathematically. Hardiman et. al. (Hardiman, Dufresne et al., 1989) counters some of Chi’s (1981) discussion of expert categorization. Hardiman found that surface feature similarity between problems could interfere with experts’ classification of problems. Hardiman also found that novices at similar levels of education employed different types of reasoning during classification of problems. Novices that used similar reasoning for classification as experts were more proficient compared to novices of similar education. Singh (Singh, 2009) more recently continued problem categorization work by asking graduate students to categorize introductory mechanics problems based on their similarity of their solutions. Twenty one graduate students were asked to categorize 25 physics mechanics problems from their own perspective and from the perspective of the typical introductory physics students whom they were teaching. Categorizations were then compared with those made by 180 introductory physics students and seven physics faculty given the same set of problems. Overall, physics professors and graduate students were more likely to point out multiple methods for solving a problem and often created several categories for the same problem. ‘Good categories’ were those which included the primary physical principle or multiple primary physical principles. Professors outperformed the graduate students in the number of ‘good’ categorizations, but the graduate students also outperformed the introductory students. Interestingly, graduate students were very hesitant of categorizing problems from the introductory student perspective because they were unable to see how the task might be useful.

**Problem categorization linked to schema acquisition and automation.**

Chi (Chi, Bassok et al., 1989) reported on self-generated explanations of nine students working out examples of mechanics problems. These students were classified as “good” or “poor” based upon performance on the tasks given to them. The “good” students generated explanations that refined the conditions of the problems and related these conditions to physical principles represented in the text. “Poor” students did not generate sufficient self-explanations and relied heavily on examples from the text. After
the study was conducted, Chi admitted that under the right conditions, studying worked examples may be a superior way to learn as compared to practiced problem solving as it imposes less cognitive load. Students may elicit prompt schema acquisition and automation given better availability of working memory resources.

de Jong and Ferguson-Hessler (de Jong and Ferguson-Hessler, 1986) expanded upon Chi’s delineation of expert and novice problem schemata differences, arguing that having the knowledge is by itself insufficient; it must be organized in a useful manner. They determined that the problem schemata must include declarative knowledge (principles, concept, formulae), knowledge of characteristics of problem situations (recognition of pertinent conditions of a given scenario), procedural knowledge for solving problems, and strategic knowledge (planning the solution). They developed a card-sorting task to elicit all the elements of problem schemata, except strategic knowledge. Sixty five elements of knowledge were printed on cards. The 65 elements were taken from a set of 12 introductory electromagnetism problems. Each problem was constructed to have at least one element of declarative knowledge, one of procedural knowledge, and one of the characteristics of problem situations. The 65 cards were then given to two groups of students. The first group consisted of 13 students, designated ‘good problem solvers’ because they scored at or above 70% on their final examination. The second group of seven students were designated ‘poor problem solvers’ because they scored less than 30% on their final examination. Students were asked to sort the cards based on their own criteria. They were given a moment to double check their piles in the end. The good problem solvers sorted the cards by problem type, while the poor problem solvers sorted by the surface characteristics. These results support the idea that good problem solvers organize their knowledge using problem schemata.

2.1.3 Transfer of Learning

Transfer of learning is defined as an application of prior knowledge from one situation to another situation. (Singly and Anderson, 1989; Reed, 1993) Transfer of learning during problem solving occurs when prior knowledge is applied to solve a new and different problem. For this project, we look to improve students’ problem solving schema using case reuse; a strategy which relies on transfer of learning between
examples and problems. We recognized a direct relationship between students’ development of knowledge organization through principle dependent problem categorization above. Now we look extend our understanding of schemata growth by attending to literature which identifies the multiple perspectives of dynamic transfer. It is important to note that a full overview of problem solving and transfer of learning is beyond the scope of this dissertation. This section will review literature from traditional perspectives of transfer of learning, as well as the contemporary models. The section will close with the consolidation of these perspectives.

**Traditional Models**

The traditional models of transfer of learning tend to focus on the cognitive aspects of transfer. For example, Thorndike’s theory of identical elements states that transfer from one activity to another occurs only if the activities share common surface features. (Thorndike, 1906) Judd’s Theory of deep structure transfer (Judd, 1908) states that transfer depends on how much of the underlying causal principles are noticed by the learner. In both of these models, the researcher has pre-defined the knowledge students transfer between given events or activities.

In more recent traditional models, researchers like Singley and Anderson (1989) utilize information processing models such that transfer is mediated by the degree to which tasks share cognitive elements. They taught students how to use one particular text editor, and then asked them to learn another. Students learning another text editor with a higher number of procedural elements in common with the first editor, required less time to learn the second editor. They also found that the time on task across editors did not correlate with similarities in the surface features among text editors.

**Contemporary Models**

Studies in transfer of learning shifted in perspective as researchers recognized a severe lack of evidence supporting these previous ideas of transfer (Rebello, 2007). Contemporary models of transfer focus on cognitive aspects of transfer like the traditional models, but they also tend to include social and cultural environment affects. The actor-oriented perspective (Lobato, 2003) conceives transfer as the personal
construction of similarities between activities based upon how the learners see situations as similar. In this sense, Lobato like many other contemporary researchers, examines transfer from the students’ point of view instead of the researcher’s point of view. Greeno (Greeno, Moore et al., 1993) defines environmental effects on transfer in terms of a learner’s affordances and constraints during engagement of an activity and the influence they have on a different activity. Bransford and Schwartz (Bransford and Schwartz, 1999) account for social and cultural factors that affect transfer and view transfer from the learners’ perspective rather than the researchers. Most importantly, contemporary models consider transfer as an active process, not attained through passive listening.

More recently, some researchers see knowledge as not transferred but reconstructed for a new context. (Hammer, Elby et al., 2005) Researchers like diSessa (diSessa, 1988) describe knowledge in terms of grain sized elements, such that very small grain sized elements are known as phenomenological primitives. A phenomenological primitive is neither right or wrong, but can be associated with an event correctly or incorrectly. An example of this would be the phenomenological primitive, *Closer is Stronger*. If one were to say their hand will get hotter as they move it closer to a burner, this is using the *closer is stronger* p-prim correctly. If one were to say, the sun is hotter today because it is closer, this is using the *closer is stronger* p-prim incorrectly. Minstrell (Minstrell, 1992) and Hammer (Hammer, 2000) also describe knowledge in this way, referring to these small grain sizes as facets or resources, respectively. Though facets and resources vary in definition, they are all relatively small ‘microscopic’ chunks of knowledge. On the other hand, researchers like de Kleer and Brown (de Kleer and Brown, 1981) describe knowledge in terms of mental models or theories constructed through interpretation of student explanation of an event or phenomena. These models are ‘macroscopic’ in that the grain sizes are much bigger than the grain size of resources and facets. The dynamic transfer of learning model (Rebello, Zollman et al., 2005) is built upon a ‘sliding scale’ of grain size, leaving no limitation in perspective or interpretation. Resources, as defined for dynamic transfer, encompass both microscopic and macroscopic transfer. Dynamic transfer is measured, not through multiple choice measurements, but through open ended investigations of how and why students activate certain resources in certain contexts. Dynamic transfer builds upon the framework
presented by Redish (Redish, 2004) such that a variety of specific theoretical perspectives may be compared. Redish’s framework is built upon a two level system: a knowledge-structure level where association patterns (any pattern that may be drawn from a dependent relationship between resources) dominate, and a control-structure level where one can describe expectations and epistemology. From this idea, the transfer framework may be revitalized through consolidation of the traditional and contemporary ways of thinking (Rebello, 2007).

**Consolidating Contemporary and Traditional Perspectives of Transfer**

Based upon the framework presented by Redish (2004), a new framework may be proposed that consolidates both the traditional and contemporary ways of thinking about transfer so that both of these two types of transfer are valued and promoted in learning (Rebello, Cui et al., 2007).

Rebello et. al. (Rebello, Zollman et al., 2005) provides a theoretical three phase model of the transfer mechanism. Phase one of the transfer mechanism occurs when the learner is primed through external inputs to activate epistemic resources exercising executive control over mental processes. Phase two of the transfer mechanism occurs when the activated ‘executive controller’ weighs the relevance of the input data and reads out the resources to be used in the reasoning process. Phase three of the transfer mechanism occurs when the activated ‘executive controller’ activates resources previously acquired to be used in the reasoning process. The learner establishes associations between the external input and pre-existing resources by this third phase.

This three phased mechanism assumes activation of associations and readout are managed by an epistemic executive controller often influenced by the learners’ subconscious measure of various conditions in the given situation. Conditions that affect this controller include emotional, motivational, and epistemological concerns. Figure 2.1 below shows a visual representation of the three phased mechanism of transfer. The figure is taken from the proceedings of the 2005 National Association of Science Teaching Conference.
There are two kinds of associations that a learner can make while problem solving using this model. The first kind of association occurs when a learner assigns information read out from the problem to an element (resource) in his or her own prior knowledge. The second kind of association occurs when a learner reads out information from the problem and establishes a link between this read out and an element of their internal knowledge structure. This second association is more difficult for students to create than the first association, as there is no direct assignment of information from the problem to a pre-created knowledge structure.

These associations are then tied to two different transfer processes. Horizontal transfer is such that the learner reads out information provided from a problem and activates a pre-created knowledge structure. Vertical transfer is such that a learner recognizes elements of the problem that activate part of their internal knowledge structure. There is no pre-created knowledge structure that aligns with the problem in vertical transfer; only in horizontal transfer. Vertical transfer requires students to recognize limitations of the model and they must decide what read out elements are significant.

In problem solving, it is important for learners to be adept at both horizontal and vertical transfer. Schwartz and Bransford (Schwartz, Bransford et al., 2005) discuss these
ideas in terms of efficiency and innovation in transfer. Horizontal transfer requires efficiency, which involves recognizing when the existing knowledge is applicable to a given situation and applying it appropriately. Vertical transfer requires innovation, which involves realizing that your current knowledge is inadequate to solve a problem and therefore new knowledge must be created for this purpose.

Dynamic transfer of both horizontal and vertical varieties is important in problems solving for learners to evolve from novice problems solvers to adaptive experts. An adaptive expert is adept at engaging in both kinds of transfer. They recognize which kind of transfer is appropriate in a given situation and adapt accordingly. The problem solving strategies that we discuss in the next section – case-based reasoning – is supposed to help learners recognize which case in their previous experience is relevant to a new problem situation, and then apply elements of the case and its solutions selectively to the problem. In other words, case-based reasoning, when optimally applied can lead to the development of adaptive expertise. Other researchers have defined similar modes of transfer.

Salomon and Perkins (Salomon and Perkins, 1989) define two types of transfer similar to the ‘horizontal’ and ‘vertical’ transfer described by Bransford and Schwartz. Low road transfer, most similar to horizontal transfer, occurs when a new problem scenario is similar to the scenario in which original learning occurred. This allows the learner to apply pre-conceived problem solving processes ‘efficiently.’ High road transfer, most similar to vertical transfer, occurs when a new problem scenario is not so similar to the original scenario. The learner is required to reflect and abstract pertinent aspects of a previous event, experience, and/or problem to help construct a solution to their new problem. In a sense, there is a significant planning stage necessary for high road transfer. Jonassen’s (Jonassen, 2003) work aligns with these consolidated models of transfer, distinguishing between well-structured versus ill-structure problem solving. Well-structured problems which primarily involve horizontal transfer, clearly define the information and goals. Ill-structured problems which primarily involve vertical transfer, and are often underspecified, require students to make assumptions about the given problem situation affecting recall of internal processes and conceptual schema.
2.1.4 Use of Analogy

Since this projects’ focus is on case reuse, or the process of using solved examples to assist with the solution of a new problem, understanding the models of analogical reasoning are necessary to complete our understanding of dynamic transfer. Analogical reasoning is a fundamental component of cognition and is often linked to problem solving through problem comparison and assertion of reasoning (Robertson, 2001). For this section, I will discuss several theories regarding the use of analogy while solving problems.

Case-Based Reasoning

Case-based reasoning (CBR) may be generically defined as the process of solving a real-world problem based on analogies (Kolodner, 1997). CBR is not a set of procedures that carry out analogical reasoning. CBR suggests a cognitive architecture, or synthetic model of analogical reasoning, that integrates our natural reasoning skills with computational processing. In other words, once a previous case is retrieved, the solution might be adapted to solve a new problem, or several pieces from several old situations might be merged and applied to the new case. Kolodner implies that the learner must learn to extract and merge important elements from previous cases, and thus come away with something in memory that can be used to plan a new solution.

From a more abstract perspective, Fauconnier and Turners’ idea of conceptual blending (Fauconnier and Turner, 2002), or conceptual integration, suggest a process of connecting concepts to create meaning on a level below consciousness. Fauconnier and Turner believed that these ‘blends,’ or connections of concepts within our minds, describe a basic function that has developed over human existence giving humans the capacity to create meaning. These blends are similar to the learners’ process of selecting the extracted knowledge elements. There is a meaning associated with the learners’ selection, though it doesn’t necessarily have to be subconscious.

Lakoff’s theory of metaphor

George Lakoff’s (Lakoff, 1987) theory of metaphor is similar to Fauconnier and Turners’ idea of conceptual blending. The theory of metaphor describes how a familiar
situation may be used to ground understanding of an unfamiliar situation. A metaphor is a cross-domain mapping in a conceptual system. A metaphor may be created if objects in a base domain are mapped to objects in a target domain. To make this idea more clear, let’s use a metaphor that is commonly used to describe resistance in electric circuits. The base domain is a circuit wire, the target domain is a water pipe. (That is not to say this metaphor is accurate or does not cause more confusion.) In this case, a thick wire would map onto a wide pipe, a thin wire would map onto a skinny pipe, and current flowing through the wire would map onto water flowing through the pipe. Thus, when we compare the resistance in the wire to the resistance in the pipe, the thick wire will have less resistance than the thin wire, as the wide pipe has less resistance than the skinny pipe.

Table 2.1 Example of a metaphor using resistance.

<table>
<thead>
<tr>
<th>Base Domain Objects</th>
<th>Target Domain Objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thick Wire</td>
<td>Wide pipe</td>
</tr>
<tr>
<td>Thin Wire</td>
<td>Skinny Pipe</td>
</tr>
<tr>
<td>Current</td>
<td>Water</td>
</tr>
<tr>
<td>Current Resistance</td>
<td>Water Flow Resistance</td>
</tr>
<tr>
<td>Thick Wire has less current resistance than thin wire</td>
<td>Wide pipe has less water flow than skinny pipe</td>
</tr>
</tbody>
</table>

Tversky’s Contrast Model

Gentner (Gentner, 1983) introduces us to Tversky’s 1977 contrast model and builds upon it. In Tversky’s model, the similarity between A and B is greater when the size of the intersection (A intersects B) of their feature sets, is greater and the similarity is less when the size of the two complement sets (A-B) and (B-A) is greater.
Figure 2.2 Tversky’s contrast model.

Gentner’s rules for analogy: structure mapping

Tversky’s theory appears to work well for literal similarities, or sets which share many exact features, but it does not provide a good account of analogy. Gentner states that the important feature of an analogy is the overall number of shared features versus non-shared features. Rather it is the “essence of the analogy” that is most relevant. Gentner (1983) uses four preliminary assumptions to discuss the rules for analogy:

1. Domains and situations are psychologically viewed as systems of objects, object-attributes and relations between objects.

2. Knowledge is represented as propositional networks of nodes and predicates. The nodes represent concepts treated as wholes; the predicates applied to the nodes express propositions about the concepts.

3. Two essentially syntactic distinctions among predicate types will be important. The first distinction is between object attributes and relationships. This distinction can be made explicit in the predicate structure: Attributes are predicates taking one argument (ie. LARGE (x) ), and relations are predicates taking two or more arguments (ie COLLIDE (x,y) ). The second important syntactic distinction is between first-order predicates (taking objects as arguments) and second- and higher-order predicates (taking propositions as arguments). For example, if COLLIDE (x,y) and STRIKE (x,y) are first-order predicates, CAUSE [COLLIDE (x,y),STRIKE(x,y)] is a second-order predicate.
4. These representations, including the distinctions made between different kinds of predicates, are intended to reflect the way people construe a situation, rather than what is logically possible.

Gentner (1983) defines base and target relations as B and T, respectively, such that the statement, “T is like a B” holds true. Symbolically, the analogy M may be represented,

\[ M: b_i \rightarrow t_i \]

where \( i \) denotes the index of the object in the base and target. With this relationship, Gentner defines an analogy as a comparison in which relational predicates, but few or no object attributes, can be mapped from B to T.

Gentner is also careful to delineate the differences between analogies and other mappings by defining domain comparisons. When a large number of both attributes and relations are mapped, there exists literal similarity. For example, The Mazda B-Series truck is like the Ford Ranger truck (they are the same truck, manufactured on the same line, with all the same parts except the nameplates.) When a large number of relations, but few attributes, are mapped, then we have an analogy. For example, the Colorado S10 truck is like the Ford Ranger truck, but not in terms of the number of shared parts. The S10 is approximately the same size, and the parts and performance are comparable to the Ford version, but there are some aesthetic attributes that remain different.
Figure 2.3 Structure map of Rutherford analogy.

Gentner herself refers to the Rutherford analogy: The atom is like the solar system. Gentner also represents this analogy using a visual representation of a structure map. See Figure 2.3 above for the Rutherford analogy structure map representation given in Gentner’s paper. Finally, when a base domain is an abstract relational structure we have an abstraction. For example, the Ford Ranger is a truck.

We also should note that Gentner’s structure mapping theory is closely paralleled by multiconstraint theory (Holyoak and Thagard, 1989), which affirms the existence of
three constraints on the use of analogies: similarity, structure, and purpose. Assuming that an analogy is a simple mapping between separate domains, Holyoak and Thagard suggest similarities in the structure of two domains will place constraints on the number of possible analogical mappings. The user must also constrain the selection of mappings by their own understanding of the purpose behind the analogy.

_Podolofsky’s Model of Analogical Scaffolding_

Podolofsky (Podolefsky, 2008) presented his own attempt to unify theories of analogy focusing on predicting how students learn through analogy. Podolefsky’s model of analogical scaffolding redefines an analogy as a mapping from a base domain to a target domain, including bidirectional projections as well as multi-layered analogies. Analogical Scaffolding does not require “stable and coherent knowledge structures that exist a priori, but allows for smaller scale schemata to be cued and blended with other schemata on the fly.”

Analogical reasoning is recognized as a compelling operation employed while problem solving. It is not, however, well-known as to how one might facilitate productive analogical reasoning. Here I present theory recognizing case reuse as a possible strategy for eliciting explicit analogical reasoning during problem solving.

_Case Reuse_

In scientific computing, case reuse refers to design patterns, where commonly occurring problems in software design may be solved with the reuse of particular patterns relating the interaction between classes and objects. (von Mayrhauser, 1994) In cognitive psychology, case reuse has been defined as the process of solving problems based on analogy. (Faltings, 1997) More recently, Jonassen (Jonassen, 2006) presents case reuse as a strategy, presenting problem solving cases as examples or analogs of how similar problems are solved. Students construct schema, or mental representations, by analyzing a worked example. This schema may be retrieved as learners work solutions to new, similar problems. These schema consist of knowledge about problem type, structural elements (acceleration, velocity, distance, etc), situations in which the problem occurs (car, on an inclined plane, baseball, etc), and the processing operations required to solve
the problem. Jonassen (2006) also reports research by Catrambone & Holyoak (Catrambone and Holyoak, 1989) and Reed (Reed, 1987) which suggests learners fail to recall examples or schema appropriately because their retrieval is based upon similarity of objects between examples, not their structural features. Catrambone & Holyoak also suggest that generalization improves when problems emphasize structural features shared with a similar example, and the number of examples is increased (i.e., three examples are better than two). Jonassen also notes that a strategy or strategies may be necessary to provide scaffolding for strengthening students’ reflection of the conditions associated with physical concepts and principles. This overlaps with Chi’s arguments for working examples where such scaffolding strategies are used to lessen students’ cognitive load, eliciting students to communicate the hierarchy of physical conditions identified as pertinent to a given solution.

Ward and Sweller (Ward and Sweller, 1990) also discuss the value of worked examples. The authors believe that problem solving skill is made up of a set of schemas and rules. The schemas, or mental representations of knowledge elements and their associations, enable students to recognize problems through those associations. The rules include automation of mathematical procedure or physical principles as equations. Given worked examples, the authors contend that students should reduce the cognitive load imposed by the search for equations and the means-end strategy. Investigations in geometric optics showed favorable performance for the cohort of students given worked examples over homework problems. However, further study using linear motion and projectile motion problems showed no evidence of effect on linear motion problems and a slightly negative effect on projectile motion problems. It was ascertained that these results were likely due to poor construction of worked examples. Students spent too much time trying to coherently compose a mental representation of the full problem using the disintegrated verbal descriptions, diagrams, and equations.

### 2.2 Research based Problem Solving Strategies

For our study, we intended to facilitate case reuse using several promising strategies. These strategies are based on research and vary widely with respect to their representation and specific intentions.
2.2.1 Introduction to visual structure maps

In our first introduction to Gentner’s Structure mapping theory, she introduces us to a structure map as a mental representation of knowledge as propositional networks of nodes and predicates. The nodes represent concepts treated as wholes and the predicates applied to the nodes express propositions about the concepts. These networks include distinctions between object attributes and relationships, as well as higher order distinctions between predicates.

Gick and Holyoak’s Radiation/Army analogy

Research that followed from this work suggests experts often recognize and use analogies spontaneously to assist in solving problems as well as build upon their own structure maps (Clement, 1998). However, Gick and Holyoak looked to determine how adept students that are not experts in their field of study might readily generate analogies. The two researchers conducted several studies using the same analogous story and problem set. The story, which I will refer to as the fortress story, was about a General’s army attacking a fortress from one direction. Because the General was susceptible to counter-attack from the sides, the General’s solution was to send his men along multiple lines of attack, dividing the opposition’s forces. The General’s men would then converge inside the fortress. The problem, referred to as Dunker’s radiation problem, requires a similar convergence resolution. In Dunker’s radiation problem, students are told that a patient has a tumor in a sensitive area inside their body. It has been discovered that x-rays of high intensity can destroy a cancerous tumor, but can also destroy healthy tissue. They are posed with the question, “How can the tumor be subjected to a beam of high intensity x-rays while leaving the surrounding tissue unharmed?”

Using two separate sets of students, Gick and Holyoak presented students with Dunker’s radiation problem, and asked them to devise a solution. Only one set of these students were told the fortress story ahead of time. Only 10% of the students who were not told the fortress story were able to come up with the convergence solution. Only 30% of students that were told the fortress story were able to apply the fortress solution to Dunker’s radiation problem. If students given the story ahead of time were told to use the story, the students that came up with the convergence solution rose to 75%. Gick and
Holyoak determined that novices could in fact use analogies productively, but it required much scaffolding.

**Structure maps versus concept maps**

Jonassen (2006) determined that novice students’ structure maps were not yet robust enough to support the analogical reasoning necessary for solving problems. The structure contained too many gaps for students to bridge without some form of assistance. He determined that a visual representation of an experts’ structure map may be an appropriate means by which students may reflect upon their own mental structure maps during problem solving. Structure maps were then redefined as visual representations of concepts and principles and their associations.

During the early development of visual maps at Cornell University by Novak (Novak, Gowin et al., 1983), concept maps were used to increase meaningful learning in science education. Novak’s maps were best described as “a visual representation expressing functional interdependency between concepts and quantities.” They may also be described as a set of propositions such that a proposition consists of two concepts and their relation to one another. Novak’s concept maps may or may not be cyclical in nature, and do include multiple conjoining propositions.

Concept maps or structure maps were widely used for problem solving. (Gentner, 1983; Novak, Gowin et al., 1983; Birney, Fogarty et al., 2005) Novak (1983) used hierarchical structure maps and knowledge Vee maps with several classes of seventh and eighth grade science students. His goal was to learn how students adapted to using maps while in the classroom and how student performance differed after the maps were integrated into their classroom activities. Vee maps are quite different from concept maps. They are visual representations in the form of a V, such that the left side usually represents conceptual knowledge and things that we know about a situation, while the right side represents the inquiry of interest and the methodology required to gain a suitable solution or answer to our initial inquiry. Novak began the year with instruction on creation and use of concept maps and Vee maps. Students were then asked to use the concept maps and Vee maps in the classroom for the remainder of the school year. Students participants present full range of abilities and were sampled from two separate
school populations. The frequency of use of concept mapping and Vee mapping in the classroom was dependent on the three teachers participating in the study, particularly the amount of class time they allotted per week for map use. Data from the study suggested that students acquired high competency in the use of the Vee heuristic and concept mapping with time (approx. 6 months), and students performed reasonably successfully for both strategies during this year long study. Data also suggested that students of all ability levels, generalized using standard measures like the SAT Examination, could be successful in concept mapping.

**Pre-created versus student-created concept maps**

Nesbit & Adesope extracted 67 standardized mean difference effect sizes from 55 studies involving 5,818 participants with student education ranging from 4th grade to postsecondary level. (Nesbit and Adescope, 2006) All of these studies met specified design criteria such that students learned by constructing, modifying, or viewing concept maps. These studies included multiple domains such as science, psychology, statistics, and nursing. Studies were selected for extraction based upon five criterion: (a) researchers contrast effects of map study, construction, or manipulation with effects of other learning activities; (b) researchers measured cognitive or motivational outcomes such as recall, problem-solving transfer, learning skills, attitude, etc; (c) researchers reported sufficient data to allow an estimate of standardized mean difference effect size; (d) researchers assigned participants to groups prior to differing treatments; (e) and researchers used random selection for assigning participants to groups, or used a pretest or other prior variable correlated with the outcome to control for preexisting differences among groups. The most frequent reason for rejecting a study was failure to control for prior differences among treatment groups. Results from the effect size extraction showed that the use of concept maps across several methodological features and instructional conditions was associated with increased knowledge retention. In comparison with activities such as reading passages, attending lectures, and participating in class discussions, concept mapping activities fared better in terms of attaining knowledge retention. Interestingly, the study also showed that activities using pre-created maps were more effective than other classroom activities. Though student generated maps fared
better as compared to the pre-created maps with respect to students overall examination gain scores, the students’ participating in these studies required notably longer training.

For our study, structure maps are pre-created maps which visually represented the associations between quantities of a given concept and/or set of principles. Six maps were created by two to four physics experts covering a range of first semester introductory physics topics. An example of a structure map used in the first phase of this project may be seen in figure 2.4 below. The structure map represents the associations between quantities of work and energy.

**Figure 2.4 Work and energy structure map.**

![Figure 2.4 Work and energy structure map.](image)

### 2.2.2 Questioning Strategy

*What is the Questioning Strategy?*

The title ‘Questioning Strategy’ can refer to many forms of academic questioning for many different subjects of study. Reciprocal Questioning strategies, for example, focus on peer interaction, and questions expressed between students in the classroom. King (King, 1990) breaks up the generation of questions into phases. Students are broken up individually and are asked to generate their own task specific questions. They are then divided into groups, taking turns putting questions forward among their group members. Ram (Ram, 1991) discusses knowledge goals and the questions expressed by a learner’s inadequate reasoning in the situation. In other words, students must generate their own
questions to alleviate concerns regarding how to solve given problems. The primary theme that resonates with all questioning strategies is that they aim to elicit students’ reflection on a given task. Like the structure mapping strategy, these questioning strategies aim to elicit students’ reflection of the distinctions between object attributes and relationships. Here, when I refer to the questioning strategy, I speak strictly in terms of Graesser’s (Otero and Graesser, 2001) psychological model of question asking known as PREG, and Graesser’s (Graesser and Hemphill, 1991) psychological model of answering known as QUEST.

Graesser’s (1991) psychological model of question answering referred to as QUEST, is a model for the creation of a solution or answer to a problem. It accounts for both open-ended questions and restricted-answer questions. QUEST may be best described by its four component scheme:

- It translates a question into a logical form and assigns it to a category (i.e., what, why, who, etc);
- It identifies information sources that are relevant to the question;
- It computes the subsets of nodes in each information source that furnish relevant answers to the particular question. Graesser then uses ‘conceptual graph structures’ to depict causal networks which include goals (reasons), outcomes (if/then statements), and the final events (consequences). These may align with some definitions of a concept map.
- Finally, QUEST considers pragmatic features of the communicative interaction such as the goals and common ground between answerer and question creator. Graesser’s QUEST model requires its users (students) to recognize causality; they must always depict relationships between nodes (goals and events) in terms of consequence, outcome, or reason.

The QUEST model is a precursor to the PREG model. The PREG model is defined as the psychological model of question asking, and PREG is not an acronym, but a shortened variation of the word ‘pregunta’, or ‘question’ in Spanish. The PREG model is used to predict the questions which might be posed given a particular event, contradiction, or problem. It can also be used to assist students with problem solving, but it does so from a different perspective than QUEST; by asking students to trigger their
own explanatory reasoning questions. Explanatory reasoning questions may be triggered (Graesser and McMahen, 1993) under four major conditions:

- If an unusual event occurs, people tend to ask questions about the cause and consequence of the event.
- If there is a situation of contradiction, people tend to ask questions in an effort to resolve the inconsistency.
- If there is a major obstacle to an objective, people will inquire about consequences and alternatives.
- Finally, if there are equally attractive alternatives, people ask questions that break the tie between the alternatives. These questions ask for the positives and negatives of each situation.

Graesser used these triggers to generate questions based upon the level of knowledge the question was looking to answer. What does X mean? (taxonomic), What does X look like? (sensory), What causes X? (causal), etc. Graesser then exposed students to these questions, effectively training them how to increase the quantity and quality of their questions in the classroom.

In this research project, PREG is fundamental in the creation of physics specific PREG questions. PREG may be used to direct students to reflect on all pertinent information in a problem before reasoning through the problem.

### 2.2.3 Finding Structural Similarities

Structure or Concept Mapping and PREG based questioning strategies are developed to elicit explicit identification of relational predicates and object attributes in a given event. The identification of such associations can play a large role in mathematics and science education research, but these are not the only strategies that may be useful to this study. It is important that we also take a look at strategies ‘closer to home’, or developed in the science and mathematics education research field. Below, I introduce three other strategies developed in the fields of math and physics, and intended to measure students’ organization of knowledge, familiarity with their organization, and/or understanding of underlying principles.
**Text Editing**

Text editing is a problem solving training strategy, where students are asked to look at problem statements and decide whether there is sufficient, missing, or irrelevant information provided. Text editing is defined as a measure of schematic knowledge, or students’ mental organization of knowledge elements and associations between such elements.

Low & Over, (Low and Over, 1990), performed three studies with 10th graders requiring them to classify algebraic word problems, and then generate solutions whenever possible. Classification of problems differed between two of the studies. In the first study, students were asked to text edit, or classify problems based upon sufficient, missing, or irrelevant information provided in the statement. In the second study, students were asked to group problems in terms of underlying structure and not surface detail. (Mayer, 1981) The third study asked students to categorize problems again, but with word problems changed to mathematical representations. All of these studies converged on the consensus that students’ that appropriately classified problem data by either successfully text editing or by successfully identifying structural similarities between problems, also achieved higher marks on examinations of general mathematical ability.

In another study, (Ngu, Low et al., 2002), students were asked to complete one of two problem solving tasks: text editing or conventional problem solving. Students assigned the text editing tasks were graded based upon how successfully they identified irrelevant or missing information. Students were scored on the conventional problems based upon calculations the students had shown. Overall, the scores showed a marked difference between the groups, with the conventional problem solving group vastly outperforming the text editing group. Ngu and his colleagues observed that text editing had no advantage over conventional problem solving in the domain of stoichiometry.

**Problem Posing**

Problem posing, (Mestre, 2002), requires students to pose their own question as it pertains to a given scenario. Scenarios often contain more than one constraining
principle. Findings from his study suggested that problem posing is a useful tool for probing the development and organization of knowledge.

In Mestre’s paper, he describes a study in which a set of four calculus-based engineering students were recruited to pose problems for three concept scenarios. An example of a concept scenario is shown in Figure 2.5.

**Figure 2.5 Sample problem posing sheet.**

![Sample problem posing sheet](image)

The problems posed were evaluated as to whether they were solvable and how well they understood why the concepts and principles presented applied to their problem. Mestre found that problems with more complex scenarios (more constraining principles) were more difficult for subjects to pose a solvable problem. On average for all scenarios, only half of the students created solvable problems overall. Interviews with students suggested that there was only a superficial understanding of concepts present, and that students posed problems by matching one piece of the scenario at a time. Scenarios were rarely looked at holistically by the students.

**Physics Jeopardy**

Physics Jeopardy Problems (Van Heuvelen and Maloney, 1999) are designed to contain multiple representations. Van Heuvelen suggests that problems that start with equations and then lead to a word description of a process requires a deeper conceptual understanding of the language of physics. Van Heuvelen’s work is based from Maloney’s (Maloney, 1993) suggestion that students understanding of the concept of force could improve if students could interpret the process described by a force diagram. For equation based Jeopardy problems students are given a mathematical expression and asked to construct an appropriate physical situation that is consistent with that expression.
\[1.5 \times 10^{-6}N = (1.0 \times 10^{-5}C) \nu (0.0107) 0.50\]

The equation shown above is an example listed by Van Heuvelen (Van Heuvelen and Maloney, 1999). This equation suggests there could be a magnetic force exerted on a charged object moving at 30 m/s relative to a magnetic field. The student would need to consider relative directions of the magnetic field and the velocity, as well as identify the quantities represented in the equation.

For diagram and graph-based Jeopardy problems, students are given a graph or diagram and asked to construct an appropriate physical situation that is consistent with that graph or diagram. This problem is not much different from the equation jeopardy problem, but the representations are different. These types of problems can be used as instructional assessments as students cannot use formula-centered, plug-and-chug problem-solving methods. They must rely on their own ability to visualize a process that is consistent with the given information. This, in turn strengthens students ability to translate between representations.

### 2.3 Summary & Limitations

In this chapter, we have reviewed literature about the differences in reasoning skills, problem solving and transfer of learning between experts and novices. We have examined transfer from both traditional and contemporary perspectives, and also how these perspectives can be consolidated to develop a general theory of adaptive expertise in problem solving.

Research has shown that combining multiple strategies during problem solving is more beneficial than a single strategy. Case reuse looks to be a promising strategy for helping students develop adaptive expertise, and may be readily facilitated by any number of the strategies above. However, it is important to note that the previous research does not experimentally describe cases where students use analogies productively without instructor guidance. Nor do researchers describe interventions that are proven successful at promoting proper use of analogy. Thus, we set up our own research goals, to promote and assess case reuse using strategies that may be easily integrated into a classroom environment. Previous research, as presented in this chapter, often seeks to observe schemata adaptations, or meta-cognitive changes. We focused our
observations on measurement of schema development, and collected information regarding students’ perceptions of implemented strategies. Our students are the largest stakeholders so it seemed appropriate to get a sense of how useful these strategies are to the users.

In our pilot phase, we focused our preliminary data collection on students’ perceptions of the usefulness of two strategies: structure mapping and Graesser’s questions. Students assigned structure mapping were asked to complete a set of questions pertaining to a structure map on work and energy concepts and paired problems. Students assigned Graesser’s questions were asked to complete the questions with regards to the same sets of paired problems. Both strategies elicited positive responses from the individual participants, but structure maps fared better overall. Students assigned the structure mapping strategy were also less apt to misunderstand questions pertaining to their treatment.

For our first phase, we continued to look at how students perceived one of the two strategies, structure mapping, given problem sets of varying degrees of similarity and differences. Some maps were preferred over others, but overall, students felt the strategy could be effective at helping them solve problems. Because students continued to show no evidence of schema development over a semester long treatment, the second and third phase of this project took a slightly different direction.

Using analogical reasoning arguments for simple comparison and contrasting of cases, a protocol was designed such that worked examples were introduced alongside unsolved problems. Step-by-step guides of problem solving included active reflection of principles involved and similarities and differences between the worked example and the unsolved problem. Students given different unsolved problems were also asked to compare and contrast their cases, and eventually pose their own problem which incorporated elements from all problems seen during the treatment for that week.

As this project has adapted, it continues to use this literature review as the foundation. These strategies used by this project include worked examples, active reflection of contrasting cases, emphasis on deep-structure elements within problem sets, and assessing students’ development of problem solving schemata using non-traditional problem tasks like text editing, problem posing, and physics jeopardy.
CHAPTER 3 - PILOT STUDY

3.1 Introduction

Research and anecdotal evidence has shown that students commonly look for examples that may guide them in the process of solving an analogous problem. This approach to problem solving is more difficult for novices because they do not focus on the characteristics of the example that are most relevant to the resolution of the new problem. The skill necessary to be effective problem solvers develops with practice, but our aim is to facilitate the progress of development over time. We look to facilitate the development of this common, but powerful approach to problem solving by using research supported strategies. The problem solving strategies used in this project were chosen for their focus on organization of knowledge and their ease of accommodating case reuse. For each strategy, students must explicitly indicate quantities given in a problem statement, the associations between quantities, and how these compare with another problem. Each strategy was used in conjunction with paired problems of similar physical principle. The two strategies were never used together. Our objective was to determine whether treatments conducted only once as extra-credit tasks, each using separate problem solving strategies, would affect student perception of problem solving strategies and/or implicitly affect student performance on solving concept-related problems.

In this chapter, I will explain these strategies and describe our methodology. Observations and results stemming from the individual interviews and examinations will also be presented. I will conclude this chapter with implications resulting from the pilot and how these affect future stages of this project.

3.2 Creation of Materials

I, along with two other investigators on this project, generated two separate problem solving strategies which focus on representations of learners’ knowledge. Each strategy was used alone as a one time treatment. We chose Work and Energy as the concepts to be covered in the treatments for several reasons. This pilot study uses a set of one-time treatments, so the concepts covered must not be so complicated by
mathematical difficulties or common underlying experiences that any effects due to the treatments are easily surmountable. Work and energy concepts were chosen because they are represented in terms of all scalar quantities. Concepts, such as force or momentum for example, would be much more complicated due to the use of vector notation needed to explain these concepts. Our choice of work and energy was also because the principles involved in that topic can be extended to other topics covered. Work and energy are concepts that cross over to other domains covered in the same semester, e.g. Bernoulli’s principle. The placement of work and energy in the timeline of the semester is also convenient. It leaves ample time at the beginning of the semester for planning and recruiting of volunteers. There is also time at the end of the semester for assessment as to whether the strategies have had an impact on student problem solving in other areas.

We implemented the two selected strategies as two separate extra-credit treatments to facilitate case reuse in problem solving by pairing problems and asking for explicit comparison. We grouped problems from algebra based textbooks based upon physical concepts required to solve the problem. Problem pairs intentionally did not share facial features. For example, a problem involving an arrow leaving a bow may use the same physical concepts as a problem involving a baseball being thrown into a mitt, but the arrow shot through a bow is not facially similar to the baseball thrown into a mitt. Three sets of pairs were created covering the topics of potential energy, work-kinetic energy theorem, and conservation of energy. When offered to students, the problem pairs always appeared in the order described above. The problem pairs were then combined with one of two treatment strategies: the questioning strategy or the structure mapping strategy.

The questioning strategy was based upon Graesser’s questioning strategy template, a generated question list that solicits students to openly communicate information relevant to the question resolution. (Otero and Graesser, 2001) It trains students to trigger questions with each problem that look to extract the interdependent relationships of given information as it pertains to a described event. It is not intended to force a particular process of resolution, but to incorporate quality questioning in students’ problem solving framework.
We produced our own physics oriented template that asked students several questions pertaining to the two problem statements shown side by side. Students were asked to identify the principles or associations between quantities, the inferences that should be drawn from the statements, and relationships between quantities given hypothetical changes in our problem scenarios. Figure 3.1 shows an example of the questioning strategy with a problem pair.

**Figure 3.1 Questioning strategy example.**

<table>
<thead>
<tr>
<th>Problem 1 (Giancoli 6-29)</th>
<th>Problem 2 (Giancoli 6-32)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 1200-kg car rolling on a horizontal surface has a speed of 18m/s when it strikes a horizontal coiled spring and is brought to rest in a distance of 2.2m. Neglecting friction, what is the stiffness constant of the spring?</td>
<td>A spring of stiffness constant 53N/m hangs vertically so that the lower end of the spring is 0.15m above the ground. A 2.5-kg mass is now attached to the spring. Neglecting air resistance, how far above the ground is the lower end of the spring?</td>
</tr>
</tbody>
</table>

Q1-1 In general, Problem 1 could be solved by applying. Select all that apply.
- a.) Newton's Second Law of Motion
- b.) Work - Energy Theorem.
- c.) Conservation of Mechanical Energy.
- d.) Conservation of Linear Momentum.

Q2-1 In general, Problem 2 could be solved by applying. Select all that apply.
- a.) Newton's Second Law of Motion.
- b.) Work - Energy Theorem.
- c.) Conservation of Mechanical Energy.
- d.) Conservation of Linear Momentum.

Q1-2 Which of the following quantities are directly given in the Problem 1? Select all that apply.
- a.) Initial speed of the car.
- b.) Final speed of the car.
- c.) Mass of the car.
- d.) Stiffness constant of spring.
- e.) Compression in spring.
- f.) None of the above. The correct answer(s) is are ________
- g.) Additional answer(s) is are ________

Q2-2 Which of the following quantities are directly given in the Problem 2? Select all that apply.
- a.) Initial velocity of the mass.
- b.) Final velocity of the mass.
- c.) Value of the mass.
- d.) Stiffness constant of spring.
- e.) Extension of spring.
- f.) None of the above. The correct answer(s) is are ________
- g.) Additional answer(s) is are ________

Q1-3 Which of the following physical quantities change in Problem 1? Select all that apply.
- a.) Kinetic Energy of the car.
- b.) Elastic Potential Energy of the car.
- c.) Gravitational Potential Energy of the car.
- d.) None of the above. The correct answer(s) is are ________
- e.) Additional answer(s) is are ________

Q2-3 Which of the following physical quantities change in Problem 2? Select all that apply.
- a.) Kinetic Energy of the mass.
- b.) Elastic Potential Energy of the mass.
- c.) Gravitational Potential Energy of the mass.
- d.) None of the above. The correct answer(s) is are ________
- e.) Additional answer(s) is are ________
<table>
<thead>
<tr>
<th>Problem 1 (Giancoli 6-29)</th>
<th>Problem 2 (Giancoli 6-32)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 1200-kg car rolling on a horizontal surface has a speed of 18 m/s when it strikes a horizontal coiled spring and is brought to rest in a distance of 2.2 m. Neglecting friction, what is the stiffness constant of the spring?</td>
<td>A spring of stiffness constant 53 N/m hangs vertically so that the lower end of the spring is 0.15 m above the ground. A 2.5-kg mass is now attached to the spring. Neglecting air resistance, how far above the ground is the lower end of the spring?</td>
</tr>
</tbody>
</table>

**Q1-4 Identify the non-conservative forces acting on the car in Problem 1. Select all that apply.**

a.) Force of the spring  
b.) Force of gravity  
c.) There are no non-conservative forces acting on car.  
d.) None of the above. The correct answer(s) is are ____________  
Additional answer(s) is are ____________

**Q2-4 Identify the non-conservative forces acting on the mass in Problem 2. Select all that apply.**

a.) Force of the spring.  
b.) Force of gravity.  
c.) There are no non-conservative forces acting on softball.  
d.) None of the above. The correct answer(s) is are ____________  
Additional answer(s) is are ____________

**Q1-5 Which of the following information that is provided in the Problem 1 is not required to solve the problem? Select all that apply.**

a.) Mass of the car.  
b.) Initial speed of the car.  
c.) Spring compression.  
d.) All information given is required.  
e.) None of the above. The correct answer(s) is are ____________  
f.) Additional answer(s) is are ____________

**Q2-5 Which of the following information that is provided in the Problem 2 is not required to solve the problem? Select all that apply.**

a.) Value of the mass.  
b.) Stiffness constant of spring.  
c.) Initial position of lower end of spring.  
d.) All information given is required.  
e.) None of the above. The correct answer(s) is are ____________  
f.) Additional answer(s) is are ____________
### Problem 1 (Giancoli 6-29)
A 1200-kg car rolling on a horizontal surface has a speed of 18 m/s when it strikes a horizontal coiled spring and is brought to rest in a distance of 2.2 m. Neglecting friction, what is the stiffness constant of the spring?

### Problem 2 (Giancoli 6-32)
A spring of stiffness constant 53 N/m hangs vertically so that the lower end of the spring is 0.15 m above the ground. A 2.5-kg mass is now attached to the spring. Neglecting air resistance, how far above the ground is the lower end of the spring?

### Q1-6
Problem 1 is changed such that rather than roll on a horizontal surface the car rolls down hill before striking the spring. What is the minimum amount of additional information you would need to find the height of the hill? Select all that apply.
- a) The stiffness constant of the spring.
- b) Angle of the incline of the hill.
- c) Speed of the car at the top of the hill.
- d) Speed of the car at the bottom of the hill.
- e) None of the above. You already have sufficient information.
- f) None of the above. Other information needed is ________________.
- g) Additional answer(s) is are ________________.

### Q2-6
Problem 2 is changed such that rather than hang in air, when the mass is attached the spring extends into a beaker of fluid. What is the minimum amount of additional information you would need to find the force of resistance provided by the fluid? Select all that apply.
- a) Distance lower end of the spring extends in fluid.
- b) Distance lower end of the spring extends before fluid.
- c) Work done by the fluid on the mass.
- d) None of the above. You already have sufficient information.
- e) None of the above. Other information needed is ________________.
- f) Additional answer(s) is are ________________.

### Q1-7
You have already solved Problem 1 and are asked to add an additional part (ii) to Problem 1 so that it is most similar to Problem 2. Write your own part II in this box and briefly explain why part (ii) makes Problem 1 resemble Problem 2.

### Q2-7
You have already solved problem 2 and are asked to add an additional part (ii) to Problem 2 so that it is most similar to Problem 1. Write your own part II in this box and briefly explain why part (ii) makes Problem 2 resemble Problem 1.

### Q1-8
Imagine that the distance of compression of the spring is greater than that given in the original problem. What is likely to cause such an increase? Select all that apply.
- a.) Mass of the car is increased.
- b.) The surface is tipped at a downward angle.
- c.) The car starts off at a distance much greater from the spring than in the original problem.
- d.) The spring constant has changed.
- e.) Friction is no longer negligible.

### Q2-8
Imagine that the spring bounces off the ground and is unable to stretch to its full length. What is likely to cause such an increase in displacement of the spring? Select all that apply.
- a.) Mass increased.
- b.) The spring constant decreases.
- c.) Air resistance is no longer negligible.
- d.) The spring is lowered.
- e.) The spring is too close to the wall and is experiencing a frictional force.

For the second treatment, students were introduced to a pre-existing visual representation of the associations of quantities relevant in Work and Energy. This visual representation is referred to as a structure map. (Jonassen, 2003) This is different from...
Gentner’s structure maps primarily because the map exists as a physical visual
representation. (Gentner, 1983) Gentner’s structure maps were internal cognitive
frameworks. We initially considered asking students to create the map on their own, but
after a careful review of literature we determined that it would be more efficient for the
researchers to create the structure map representing all of the major quantities used in
work and energy problems. Given the limited access to students and time available, it
would not be feasible for us to ask students to create maps representing their own
structure of knowledge without long-term training. It has also been shown by Nesbit &
Adescope (2006) in a meta-analysis of concept mapping studies that pre-created maps by
experts can be effective in achieving performance gains.

For our treatment, students were presented with the structure map and a problem.
Unlike our questioning strategy, the problems in a pair were not shown side by side.
Both treatments were introduced to students via an internet interface created by our
collaborators at the University of Missouri. The designers felt it would be too difficult to
create a page that would be clearly visible on the computer screen which included
problems side by side for the structure mapping treatment. Instead individual problems
were shown underneath a .jpeg image of the structure map. Problems in a given pair
were always seen consecutively and the sets were given in the same order as that seen by
the questioning strategy. Students were required to answer a set of questions below each
problem by selecting the number that corresponded to the node they wanted to identify.
The questions asked students to identify quantities given, asked for, and not needed in the
problem statement. Students were also asked to identify quantities in the map that were
necessary in the solution, but were not apart of the quantities given in the statement.

Figure 3.2 shows the structure map that was provided to the students and an
example of questions that were provided to the students for use with the structure map.
Each column of questions refers to a different problem in the problem pair.
Figure 3.2 Structure map and questions for work-energy problem pairs.

### Problem 1
A 0.088kg arrow is fired from a bow whose string exerts an average force of 110N over a distance of 0.78m. Neglecting air resistance, what is the speed of the arrow as it leaves the bow?

**Q1-1** Check all quantities that are directly given in Problem 1. ([List of check boxes with # 1 through # 23 on the structure map above](#))

**Q1-2** Check all quantities that are not directly given in Problem 1, but you are expected to know so that you can solve the problem. ([List of check boxes with # 1 through # 23 on the structure map above](#))

**Q1-3** Check all quantities that you are asked to find in Problem 1 ([List of check boxes with # 1 through # 23 on the structure map above](#))

### Problem 2
A 0.25kg softball is pitched at 26m/s. By the time it reaches the plate a distance 15m away it has slowed to 23m/s. Neglecting gravity, what is the average force of air resistance during the pitch?

**Q2-1** Check all quantities that are directly given in Problem 2. ([List of check boxes with # 1 through # 23 on the structure map above](#))

**Q2-2** Check all quantities that are not directly given in Problem 2, but you are expected to know so that you can solve the problem. ([List of check boxes with # 1 through # 23 on the structure map above](#))

**Q2-3** Check all quantities that you are asked to find in Problem 2 ([List of check boxes with # 1 through # 23 on the structure map above](#))
### Problem 1
A 0.088kg arrow is fired from a bow whose string exerts an average force of 110N over a distance of 0.78m. Neglecting air resistance, what is the speed of the arrow as it leaves the bow?

### Problem 2
A 0.25kg softball is pitched at 26m/s. By the time it reaches the plate a distance 15m away it has slowed to 23m/s. Neglecting gravity, what is the average force of air resistance during the pitch?

### Q1-4 Check all quantities that you need to calculate as you are solving Problem 1. (These are the intermediate quantities – neither those that are given, nor those that you are asked to find) (List of check boxes with # 1 through # 23 on the structure map above)

### Q2-4 Check all quantities that you need to calculate as you are solving Problem 2. (These are the intermediate quantities – neither those that are given, nor those that you are asked to find) (List of check boxes with # 1 through # 23 on the structure map above)

### Q1-5 Check all quantities given in Problem 1 that are not required for solving the problem. (Leave quantities unchecked if you decide none of the given quantities are unnecessary i.e. all of the quantities that are given are required to solve the problem.) (List of check boxes with # 1 through # 23 on the structure map above)

### Q2-5 Check all quantities given in Problem 2 that are not required for solving the problem. (Leave quantities unchecked if you decide none of the given quantities are unnecessary i.e. all of the quantities that are given are required to solve the problem.) (List of check boxes with # 1 through # 23 on the structure map above)

### 3.3 Extra Credit Implementation

Student participants’ were asked to access the treatments online as an extra credit activity in the class in which they were enrolled. As students logged into the system, they were randomly assigned to one of three groups: questioning group, structure mapping group, or control group. The questioning and structure mapping groups each worked out three principle types of problem pairs that were based on the following principles: work-energy theorem problems, potential energy problems, and conservation of energy problems. After completing these tasks they were given a set of three problems to print out and solve on a sheet of paper showing their complete solution, and hand in to the instructor.

The control group was given one problem of each principle type, instead of a pair of problems, and they were asked to explicitly solve each of the problems on a separate piece of paper showing their complete solution. This task ensured that the control group had the same time on task as the treatment groups. Students in the treatment groups (questioning or structure mapping) were not explicitly asked to solve the problems given in the online activity, but to answer questions that were central to the strategies. All three
groups (questioning, structure mapping and control) were also given a set of three problems that each were dependent on one of the three principle types (potential energy, kinetic-energy theorem, conservation of energy), but were not the same as any of the problems seen in the treatment. We asked students to solve each of these problems on a separate piece of paper.

### 3.4 Results - Extra Credit Assignment

Students were randomly assigned one of three groups: questioning, structure mapping or control. Each group worked with three types of problem pairs: work-energy theorem problems, potential energy problems and conservation of energy problems. After students assigned to the questioning or structure mapping groups completed the three problem pairs, they were given three different problems to solve and hand in, one from each type. The control group was asked to solve and hand in six problems, two of each type. The structure mapping and questioning groups were assessed based on completion and correctness of answers given on the questions provided underneath problems in the online portion of the extra-credit assignment. Students were not asked to solve the problems given, but they were asked to answer questions related to the three distinct problem pairs. Master solutions sets were created for both treatments’ question sets and student answers were graded on correctness. It was important for the researchers and primary instructor to get some feedback as to whether students were taking the assignment seriously, and to determine how students responses compared with expert responses. Means and standard deviations were calculated for both groups. Pearson correlation coefficients (two-tailed) were calculated between the different problem pair sets in each treatment to determine whether there existed a correlation between student correctness between pairs.

Assuming the correlation is significant at the 0.01 level, the structure mapping group showed strong correlation between all sets. The questioning group showed strong correlation at the 0.01 level between work-energy pair scores and conservation of energy pair scores, and between potential energy and conservation of energy pair scores. Work energy and conservation of energy pair scores would be significantly correlated at the 0.05 level, but work energy and potential energy problem pair scores showed no
significant correlation with a correlation coefficient of 0.125. See the tables 3.1 and 3.2 below for a summary of the results described above, or see Appendix A-1 for complete statistical overview of the treatments.

Table 3.1 The questioning strategy group Pearson correlation values.

<table>
<thead>
<tr>
<th></th>
<th>Set 1 Corr./(Sig.)</th>
<th>Set 2 Corr./(Sig.)</th>
<th>Set 3 Corr./(Sig.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set 1-Potential Energy</td>
<td>1/-</td>
<td>.375/(.125)</td>
<td>.471/(.049)</td>
</tr>
<tr>
<td>Set 2- Work-Kinetic Energy Theorem</td>
<td>.375/(.125)</td>
<td>1/-</td>
<td>.797/(.000)</td>
</tr>
<tr>
<td>Set 3-Conservation of Energy</td>
<td>.471/(.049)</td>
<td>.797/(.000)</td>
<td>1/-</td>
</tr>
</tbody>
</table>

The significant correlation between sets, with few exceptions, is a good sign that the overall performance between pairs remains similar, and that students are not showing significant signs of fatigue in the final pair set.

The number of questions asked between treatments was not the same, so the results presented on student mean scores is reported in terms of mean percentages and percent standard deviation. The mean percentage correct for the structure mapping group was 51.2%, with a percent standard deviation of 22.6%. The mean percentage correct for the questioning group was 56.5%, with a percent standard deviation of 10.0%. The standard deviation for our structure mapping group, 22.6%, resulted from a much larger variance in answers to questions asked than the questioning group.

Table 3.2 The structure mapping group Pearson correlation values.

<table>
<thead>
<tr>
<th></th>
<th>Set 1 Corr./(Sig.)</th>
<th>Set 2 Corr./(Sig.)</th>
<th>Set 3 Corr./(Sig.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set 1-Potential Energy</td>
<td>1/-</td>
<td>.613/(.000)</td>
<td>.556/(.002)</td>
</tr>
<tr>
<td>Set 2- Work-Kinetic Energy Theorem</td>
<td>.613/(.000)</td>
<td>1/-</td>
<td>.940/(.000)</td>
</tr>
<tr>
<td>Set 3-Conservation of Energy</td>
<td>.556/(.002)</td>
<td>.940/(.000)</td>
<td>1/-</td>
</tr>
</tbody>
</table>

Table 3.3 The mean percentage correct on questions given during the structure mapping and questioning strategy treatments.

<table>
<thead>
<tr>
<th>Treatment Questions</th>
<th>Mean Percentage Correct</th>
<th>Percent Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questioning Group</td>
<td>56.5%</td>
<td>10.0%</td>
</tr>
<tr>
<td>Structure Mapping Group</td>
<td>51.2%</td>
<td>22.6%</td>
</tr>
</tbody>
</table>

Students’ solutions to the three problem types were compared with regard to both completion and correctness. Students in the control group were graded on all six
problems handed in. Students in the treatment groups were graded on the three problems
given directly after completing the treatment tasks.

Each problem was graded out of five points for a total of 15 points for the
treatment students and 30 points for the control group. Four (4) out of the five (5) points
were given independent of whether the final solution was correct. A student scored one
(1) point if they wrote down the correct principles and/or applicable equations related to
those principles. One (1) to three (3) points were assigned to a solution dependent upon
the extent which a student logically applied a given equation/formula to arrive at a
solution. Students were given some or all of these points if they were able to correctly
associate some or all of the quantities given in the problem statement with the quantities
in the equation. Five (5) points were given if the solution was fully correct. All
problems were graded by the same grader.

Table 3.4 Grading rubric for all problems given as part of the extra-credit
assignment.

<table>
<thead>
<tr>
<th>Description of solution displayed</th>
<th>Points allotted (out of 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(points should not be rewarded if student lists multiple irrelevant principles or equations with respect to the solution)</td>
<td></td>
</tr>
<tr>
<td>Correct principle stated or implied by equation selection</td>
<td>1 pt</td>
</tr>
<tr>
<td>Logical application of equations/ formulas necessary for solution</td>
<td>2-3 pts</td>
</tr>
<tr>
<td>Correct Answer (no mathematical error)</td>
<td>1 pt</td>
</tr>
</tbody>
</table>

The control group, which completed six extra credit problems, was scored out of
30 points. The treatment group, which completed three extra credit problems, were
scored out of 15 points each. Due to a difference in possible maximum score, the data
were translated to percentage scores. Means, standard deviations, and Pearson
correlation coefficients were calculated to determine whether there were substantial
differences between student performance on the problems or between types of problems.
Students in the structure mapping treatment had an average score of 69.0% with a 15.0%
standard deviation. Students in the questioning treatment had an average score of 72.4%
with a 10.1% standard deviation. Students in the control group had an average score of
72.0% with a 12.4% standard deviation. Student participants showed no significant difference between the three treatment groups on the extra credit problems.

**Table 3.5 Student performance on problems assigned with extra-credit.**

<table>
<thead>
<tr>
<th>Problem performance</th>
<th>Mean Percentage Correct</th>
<th>Percent Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questioning Group</td>
<td>72.4%</td>
<td>10.1%</td>
</tr>
<tr>
<td>Structure Mapping Group</td>
<td>69.0%</td>
<td>15.0%</td>
</tr>
<tr>
<td>Control Group</td>
<td>72.0%</td>
<td>12.4%</td>
</tr>
</tbody>
</table>

Two one way ANOVA tests were conducted (See Table 3.5 and Table 3.6): One on the problem percentage scores of student participants and the other on the final student score (sum of the treatment question score and the problem score, or just the problem score for the control group). Both ANOVA tests showed no statistically significant difference between the three treatment groups on the extra credit problems or on the resulting final scores.

**Table 3.6 ANOVA problem performance significance between groups.**

<table>
<thead>
<tr>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>127.377</td>
<td>2</td>
<td>63.689</td>
<td>.381</td>
</tr>
<tr>
<td>Within Groups</td>
<td>8848.615</td>
<td>53</td>
<td>166.955</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>8975.992</td>
<td>55</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 3.7 ANOVA calculation of final score significance between the three groups.**

<table>
<thead>
<tr>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>.565</td>
<td>2</td>
<td>.283</td>
<td>.086</td>
</tr>
<tr>
<td>Within Groups</td>
<td>307.268</td>
<td>93</td>
<td>3.304</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>307.833</td>
<td>95</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 3.5 Far Transfer Problem (Examination) Implementation

The extra-credit assignment given to all students was based upon the concepts of work and energy. With permission from the primary instructor, we were allowed to create a problem to be asked on their exam covering work and energy (See Figure 3.3).
The *Descriptive Physics* class covers, in one semester, material that is typically covered in other introductory physics courses in two semesters. Thus, our problem was the only work and energy problem on the written portion of the exam. The far transfer exam problem required the use of explicit and implicit knowledge from the problem statement and required similar mathematical progression through work and energy quantities as previous homework and extra-credit problems. This problem was referred to as a ‘far-transfer problem’ because it is presented ‘far’ after the treatment took place. The problems given directly after the treatments might be called ‘near-transfer’ problems, as they are given directly after the treatment took place. All problems given to students, directly or not-directly after treatment attempt to assess student performance with regards to completion as compared between treatment and control participants.

**Figure 3.3 Work-Energy examination problem and solution.**

3.6 Results - Examination problem

All students in the class were given a far transfer problem on their second course examination, to assess the influence of the treatment in previous extra credit exercises.
The problem, shown in figure 3.3, was graded using a rubric emphasizing students’
ability to correctly determine the appropriate equations necessary and student
understanding of the principles involved as described in their own words. (See table 3.8
for the finalized rubric.)

**Table 3.8 Rubric to assess performance on the 'far transfer' examination problem.**

<table>
<thead>
<tr>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>No answer; Draws pictures without any Energy notation, only quantities given in problem statement.</td>
<td>0</td>
</tr>
<tr>
<td>Used formula(s) describing force, velocity, or momentum without using idea of conservation</td>
<td>1</td>
</tr>
<tr>
<td>Recorded formula(s) describing energy without applying to problem solutions and without using idea of conservation</td>
<td>2</td>
</tr>
<tr>
<td>Used formula(s) describing energy without using idea of conservation</td>
<td>3</td>
</tr>
<tr>
<td>Used some application of conservation of energy</td>
<td>4</td>
</tr>
<tr>
<td>Applied Conservation of Energy and applied Work-Energy relationships. Conceptual errors occur.</td>
<td>5</td>
</tr>
<tr>
<td>Applied Conservation of Energy and applied Work-Energy relationships. Conceptually correct, only mathematical errors.</td>
<td>6</td>
</tr>
<tr>
<td>Completely correct.</td>
<td>7</td>
</tr>
</tbody>
</table>

The rubric was assessed for its inter-rater reliability (i.e. its consistency to yield the similar/same scores when used by different researchers). The rubric proved to be acceptably reliable using Pearsons’ product moment correlation coefficient, mean of difference, and limits of agreement. Pearsons’ product moment correlation is an appropriate test for reliability between raters, but it can indicate perfect correlations falsely if the variance between scores remains similar. An example of this would be when one rater always rates exactly +1 higher than the other rater. There would be a strong correlation between scores because they are all related by the same factor of +1, but this means the rubric is still unreliable. To negate this effect, the Pearson product moment correlation should be used in conjunction with another reliability test. Pearsons’ product moment correlation coefficient can be used to measure pairwise correlation among raters using a rating scale that is continuous. Our rating scale is an interval scale, and thus qualifies as continuous. The Pearson coefficient between our raters was 0.941, well-within the 0.01 level of significance. This is a very strong correlation. The mean of
difference, or average difference between rated scores among the two raters was 0.312. The confidence limits calculated at 0.05 significance level, were ±0.78. The low mean of difference of 0.312 between scores could mean that either the raters rated the problem solutions very close to one another, or the variance in the ratings was so high, that the average difference was in turn low. A confidence limit of ±0.78 demonstrates that the mean of difference was more likely low due to minimal difference between ratings as the estimated difference between raters is less than 1 at a 95% confidence.

After consultation between raters, it was determined that dissimilar ratings were due to the same factor. One rater, identified as rater 1, would identify some equations written on the paper as being ‘used’ even though they were either unused in the problem solution or they expressed a physical principle incorrectly, e.g. F= \( \frac{1}{2} kx^2 \). The other rater, rater 2, determined these equations were either not ‘used’ by students, or their physical meaning should be taken from the student’s perspective. For example, F= \( \frac{1}{2} kx^2 \), would be expressing a force, not an energy. Rater 1 determined that the dissimilarity in ratings was human error, and not a problem with the phrasing of the rubric. Rater 1 re-evaluated the problems and in turn, matched rater 2’s evaluation. The same problems were also coded by both raters on the concepts covered while solving the problem. Six coded problem solutions did not identically match. In all six cases, the majority of the codes selected to represent the concepts remained the same, though one rater would include extra codes identifying concepts included with equations written but unused in the problem solution. The other rater ignored equations that seemed unrelated to the solution.

The mean score of the structure mapping treatment groups for the examination problem was 2.71 with a standard deviation of ±1.65. (See Table 3.9) The questioning treatment groups’ mean score was 2.24 with a standard deviation of ±1.66. The control groups’ mean score was 2.73 with a standard deviation of ±1.85. There was no significant difference between solutions to the far transfer problem on the examination. The control group scored marginally, but not significantly better than the treatment group. This result is consistent with previous studies which suggest that students given a short term treatment of any problem solving strategy do not show marked improvement (Novak et al., 1983).
Table 3.9 Rubric scores for treatment and control groups on the examination far transfer problem.

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean Score</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questioning</td>
<td>2.24</td>
<td>±1.66</td>
</tr>
<tr>
<td>Structure Mapping</td>
<td>2.71</td>
<td>±1.65</td>
</tr>
<tr>
<td>Control</td>
<td>2.73</td>
<td>±1.85</td>
</tr>
</tbody>
</table>

3.7 Interview Implementation

The pilot study also involved two sessions of semi-structured interviews with eight students. The students were a stratified sampling of volunteers. After a complete list of volunteers was collected from a visit to the lecture, students were categorized by their extra-credit treatment/control participation. Volunteers from each group were randomly selected and contacted to participate in individual interviews. Few volunteers responded to the email request, however, eight volunteers agreed to participate in both interviews. Two students were a part of the questioning strategy group, two others were apart of the concept mapping group, and the remaining four were apart of the control group. Both interviews followed the extra credit and the second in-class examination. The students were not made aware that the interviews would relate to the extra-credit assignments.

We began the first interview for all eight students by asking them to solve a problem involving work and energy (See Figure 3.4). They were not told that the problem required the application of work or energy, but they were given a copy of their course textbook and a calculator to use during the interview. After each student completed their attempt to solve the problem, we asked students questions about their work. Questions asked during the first interview focused on the mathematical procedure provided by the student, the strategies used while working the problem, and difficulties students articulated as they explained their work. See Appendix A-3 for Interview 1 protocol questions. Students were asked to return for a second interview one or two weeks after the completion of the first interview.
While the purpose of the first interview was to gain insight on students’ strategies for solving problems, the second interview was focused on acquiring information that could potentially validate the group strategies and/or their goals. The second interview required separate protocols for each group. We first asked students in both the structure mapping and the questioning groups to recall their extra-credit assignment as best they could. This question was important since it was a little more than a month since students had completed the extra credit assignment. If students were able to recall any part of the extra-credit assignment, they were asked to apply what they remembered to the problem they were provided during the first interview. If they were unable to recall the extra credit assignment, they were then provided a copy of the extra-credit treatment and asked to re-examine it. After the participants reported that they had re-acquainted themselves with the treatment they had been previously provided in the extra credit assignment, we concentrated our interview questions on the students’ views about the intended purpose behind the structure mapping or questioning strategy. If time allowed, students were asked to use their extra-credit strategy while attempting for a second time to solve the
problem given in the first interview. See Appendix A-4 for Interview 2 protocol for the structure mapping group or questioning strategy group.

Students in the control group were asked questions similar to students in the treatment groups during their interviews. Since these students had not been exposed to either of the treatments in the online extra credit assignments, we did not ask them to review the extra-credit assignment. Rather, we used this time to explain to them one of the two strategies – questioning or structure mapping that was provided to the treatment groups in the online extra credit. Each control group student was then asked to work through a portion of one of the other extra-credit assignments. This task allowed them to get acquainted with the strategy they had just been explained to them. From that point onward, the interview protocol was identical to that used for students in the structure mapping or questioning groups. See Appendix A-4 for Interview 2 protocol for the control group.

At the end of each interview, handouts were given to all students, asking them to spend some time reflecting on their own thinking about how they might approach solving the given problem and whether they might gain insight from the problem solving strategy. The handout gave students the opportunity to candidly offer their opinion of the extra-credit assignment and what aspects of the assignment they found to been useful, not useful, or confusing while attempting to solve the problem from the first interview.

3.8 Results - Interviews

The second half of this study involved two 50-minute sessions of semi-structured interviews for each of eight volunteers. Student volunteers were selected based upon the extra credit assignment they had completed. Two students each were selected from the questioning strategy group and structure mapping group, the remaining four were from the control group. The first interview, following students participation in the extra credit assignment and the second in-class examination, investigated students’ general problem solving approaches. Students were asked to work through a work-energy problem (Figure 3.4). They were allowed to use their course textbook and a calculator. After each student completed their attempt to solve the problem, we asked them questions about their solutions to the problem. The first interviews were important for gaining
insight into each of the students’ typical problem solving strategies, so that these strategies could later be compared with the strategies they used during the second interview. Students consistently used the *working-backward or means-end* approach during the first interview session, so it became apparent when students reverted back to that approach in the second interview session. (Anzai, 1991) It is important that we note that no student completed the problem given in interview 1.

Students returned for a second interview one to two weeks later. The second interview focused on acquiring information about students’ perceptions of the treatments they had completed in the extra credit assignment. Participants were initially given time to reacquaint themselves with the assignment, and we concentrated our interview questions on students’ views about the intended purpose underlying the structure mapping or questioning strategy. If time allowed, students were asked to use their strategy while attempting, for a second time, to solve the problem given in the first interview. Several themes emerged regarding the students’ perceptions of these strategies in the second interview session. These themes will be discussed in the following section. Participants were also given a worksheet at the very end of the second interview, asking each student to explicitly write out why they liked or disliked the strategy, and how it might be helpful while solving the interview problem. Unfortunately, only two students of the original 8 turned in the final worksheet. Data collected from the two worksheets were compared with statements made by the same two students during interviews. It was obvious to the interviewer that the worksheet data were not any more insightful, and there were no contradictions found between the interview discussions and the worksheet write-ups. Below we discuss results from the interviews based on the strategy introduced or reintroduced to the student.

### 3.8.1 Questioning Strategy

The questioning strategy was introduced/reintroduced to four of the eight interviewees. Two of these students belonged to the questioning group in the extra credit assignment. The other two students were from the control group and had not been previously exposed to the strategy in our study. The following themes emerged from the analysis of the interviews.
**Purpose:** All four students said that the strategy was purposeful and similar to their own problem solving techniques. When asked to explain the possible purpose of the questioning strategies, all interviewees replied similarly,

“…to help us visualize the problem...to help us think of what we should take into account, and get us thinking of what shouldn’t be taken into account…”

**Comparison with Existing Strategies:** Interviewees were also asked to explain how the questioning strategy varied from their own strategies. Responses showed that students believed the strategy was designed to mimic good question asking procedure and said that they already ask themselves the same or similar questions when they solve problems.

“Well, I always ask myself this, which is what I am given, or what’s implied in the problem. So like this, it’s talking about how far the arrow went, I’d have to take in gravity, but it’s not given in the problem, but I know what it is…”

**Use of Equations:** Responses showed question miscommunication when asked to identify from a list, concepts, laws and/or theories applicable to the problem. Three out of four responses reflected interviewees’ use of equations in the process of identification of concepts.

“Umm, on this one (question 2), I usually try to find the equation I’m using from what I’m given or I try to find an equation with a lot of what I’m given in it and try to see if there is something missing that I need....”

**Recognition of problem pairs:** Overall, students responded affirmatively when asked if they felt the questioning strategy was helpful for solving physics problems, and if they felt comfortable using the strategy. Furthermore, all students who used the questioning strategy recognized that the problems were paired in the extra credit assignment. Three of the four students articulated a reasoning for the paired problems:

“Umm...they’re both dealing with the same uh...work and energy but they’re showing it in different ways, like this one is using a spring compression in order to move the arrow and this one is just using a human just throwing it and it tells you the initial speed, but it’s still using the same type of equations.”
It is interesting to see that students perceived the questioning strategy as being useful for training them to think more explicitly about relationships in a given problem. It is also interesting to see that students, given the questioning strategy, recognized that problems are paired by principle similarities. Unfortunately, the second individual interviews provided no measurable difference between the control and questioning treatment group participants. Students’ reflections were often similar among the two distinct groups, and no one from this population was capable of solving the problem by the end of interview 2. One student from the control group completed part (a) in both interview 1 and interview 2.

### 3.8.2 Structure Mapping

The structure mapping strategy was introduced/reintroduced to the other four of the eight interviewees. Two of these students belonged to the structure mapping group in the extra credit assignment. The other two students were from the control group and had not been previously exposed to the strategy in our study.

**Purpose:** The structure map was highly regarded as a tool that assists in finding the most applicable equations. None of the four students expressed any difficulties about using the structure map. All students liked how the map represented all of the problem information. Two students, each participating in different groups, liked the way all the quantity relationships were apparent.

“...half the time its hard for me to figure out what equation to use, but like when you figure out like what it gives you and then how to figure out what equation to use from the arrows, helps, like it doesn’t give you the equation but it tells you what you need in order to figure out how to get the answer.”

**Comparison with Existing Strategies:** When asked to compare this strategy with the one that they used, all students found that the structure mapping strategy was quite different from their own problem solving technique, nevertheless they found structure mapping helpful in understanding “what you need for a problem.”

**Recognition of problem pairs:** Only one student from the structure mapping interviews recognized that the problems were paired. He found the pairings useful for
comparing question answers between the two problems. Other students, when asked if there might be a reason why the problems are paired responded, “I think there is a reason, I just don’t know what it is.”

Overall, students felt the structure map was easy to use after given the appropriate PowerPoint training slides. At the end of the second interviews, all four students worked out at least one part of the interview problem using the structure map. Unlike the questioning strategy, all four students’ solutions improved from the previous interview, as all four students were capable of successfully completing part (a) and two students (each from different groups) were capable of completing part (b) successfully. Again, there was no measurable difference between the control group and the treatment group when comparing student perceptions or interview problem performance. Students participating in the structure mapping interviews were capable of completing more parts to the interview problem on average with approximately the same amount of time spent on task.

### 3.9 Summary of Results

Students in all three groups performed similarly on all performance based assessments completed during and after the extra-credit assignment. There were no significant differences in student performance on written problem solutions given on the extra credit assignment. There were no significant differences between groups on the “far transfer” examination problem. Individual interview problem performance was slightly better for students participating in the structure mapping individual interview, but the student participants that shown improvement were equally distributed between the extra-credit control and structure mapping groups.

There are a couple reasons as to why the treatments may not have performed as well as we hoped. It is likely that the treatment was simply not long enough. Research has shown that one-time implementations are not as effective as long-term treatments. It is also possible that the treatment implementation via the web required a more user friendly interface. The one provided for these treatments could be described as somewhat difficult to navigate, particularly for the structure mapping strategy. For the interface used in this implementation, the structure map included numbers in each of the quantity
nodes such that students could identify the node by its number. Students then had to answer each of the questions by selecting the pertinent numbers directly below each question statement. This required students to scroll up and down to read the question and view the image of the structure map. It might be prudent for future investigation to be more cognizant of the user-friendliness of our interface such that problem pairs are side by side for both treatments, entertaining a better opportunity for problem comparison.

Results pertaining to problem solving performance remain inconclusive between the two treatment groups and the control group. However, this pilot study provided us with some useful insights. Our results from the individual interviews indicated that students believed these strategies are helpful, improving problem visualization and facilitating their ability to identify important information from the problem.

Students introduced to the questioning strategy believed the questioning strategy was similar to their own problem solving techniques, providing well structured questions that attempted to draw important information from the problem statement. Students introduced to structure mapping believed the structure maps were not comparable to their own problem solving resources, but still felt the strategy was an effective way to identify quantities and interpret the relationships between quantities presented in the problem.

All eight students agreed that the purpose of the strategies was to help them work out problems, though the intended purpose of some of the questions from the questioning strategy were not clear to the students. Students showed difficulty expressing differences between concepts and equations. Three of our four participants provided equations that fit some or all of the quantities given in the problem as an appropriate means of defining the problem concepts.

### 3.10 Limitations and Future Studies

The goal of the pilot study was to examine student perceptions of each strategy and assess the effectiveness of the strategies. We determined that the effectiveness of the strategies should be gauged in the long term, and that the one-shot treatment method was a better suited for gaining student insight on how their problem solving strategies related to those used in our treatments.
Based on our results above, we concluded that no significant changes needed to be initially made to the structure mapping strategy, though thought should be placed in the design of any internet interface used for future implementations. The questioning strategy would require the most alteration prior to future implementation. The intended meaning of the questions was sometimes misaligned with students’ interpretation. Questions asking students to identify concepts, theories or laws were ultimately answered with equations containing quantities identified in the problem. These questions need to be reworded in the future such that the student does not answer them in terms of an equation.

Based on these results and limitations, we concluded that our next phase of this study would include a long term quantitative and qualitative investigation of students’ performance and perception using one of these two strategies in an algebra-based physics course. It would be difficult to qualitatively assess two long-term treatments in the same semester, so structure mapping was chosen as the primary treatment method under study. It was chosen because it was well-liked by students, was an acceptable form of concept mapping as defined by the research community, and showed no weaknesses in its designed purpose. The questioning strategy was incomplete in the sense that students were unable to cite concepts or principles involved in a problem without referring to an equation.

In the following semester, we conducted a long-term implementation of the structure mapping treatment in which algebra-based students learned principles of mechanics, heat, fluid, waves, and sound. We met with two focus groups of six paid volunteers each, nine times over the semester. During each focus group learning interview, the groups practiced using structure maps with four problems over the hour. The problems differed from each other in either facial or structural features, or both. Problems were selected and modified from the chapter covering their previous homework assignment. Chapter 4 describes this next phase.
CHAPTER 4 - PHASE I - Structure Mapping Treatment

4.1 Introduction

Following the pilot study, structure maps were selected as the primary strategy used to support case reuse. A structure map is best described as a visual representation expressing functional interdependency between concepts and quantities (Gentner, 1983; Novak, Gowin et al., 1983). Gentner’s Structure Mapping Theory describes mapping as a cognitive function, or a set of interpreted implicit restraints maintained by an individual. For this project we externalize Gentner’s representation as a modified form of a concept map. By this definition, structure maps, may be categorized as a specific variation of a concept map, or a map of relationships between concepts. It was determined that two alternate treatments, such as those implemented in the pilot study, would be difficult to qualitatively assess in the long-term. The structure map was well received by our pilot participants, and no difficulties arose from use of the maps during the one-time implementation in the pilot phase. Because the maps are also a variation of a concept map, previous research results are more readily available distinguishing concept maps as an accepted strategy associated with increased problem solving performance. Research has shown that students’ use of concept maps, across several methodological features and instructional conditions was associated with increased knowledge retention (Nesbit and Adescope, 2006). Previous investigations have also reported that over the long term students can acquire procedural automation of concept mapping and assimilate it into their problem-solving repertoire (Novak, Gowin et al., 1983; Ericsson, Krampe et al., 1993).

The questioning strategy treatment used in the previous phase was also well liked by students, but opened itself up for student misinterpretation of the questions involving concepts and principles. The questions posed to students during this strategy would need to be further developed and validated for their ability to explicate deeper, thoughtful questioning by students, as the strategy was originally intended. The template used in our previous phase was significantly altered from the original template created by Graesser in order to adapt the strategy to the context of physics problem solving and case reuse.
In our previous study, students would often be unable to
distinguish a principle or concept without first referring to the equation which best
matched quantities given and asked for in the problem statement. For these reasons we
selected the structure mapping strategy over the questioning strategy for this next phase
of our project.

This chapter reports on a semester-long treatment of using structure maps in an
introductory algebra-based physics course. Our objective was to measure student
reaction to these structure maps over the semester while also observing the student use of
the maps while solving given problems. During nine focus group learning interviews, we
investigated how students performed on each of the problem sets, how students’
compared and contrasted problems of varying similarity, how the maps were perceived to
be useful while problem solving, and how the structure maps affected collaborative
problem solving. The maps evolved over the treatment in response to feedback provided
by students.

Student performance was also assessed individually on the effect of the long term
structure mapping training on students’ conceptual schema, or mental organization of
knowledge. Data were collected from two individual interviews, one at the mid-point of
the semester and the other at the end of the semester. We looked to determine whether
students participating in our focus group learning interviews were capable of progressing
through a series of non-traditional problem solving tasks. We examined the
completeness and correctness of students’ responses on problems posed in each task. We
also looked to determine how our cohort compared with the baseline group of students on
each task. Figure 4.1 below represents the research design timeline for this phase of our
project.
The following research questions are answered in this phase of the study:

- How do students use structure maps while problem solving?
- How useful do students perceive the structure maps?
- How does student performance change, if at all, over the semester using non-traditional problem solving tasks as an assessment aimed at measuring students’ efficiency of information processing?
- How does our cohort compare with the baseline group of students on non-traditional problem solving tasks?

### 4.2 Methodology – Focus Group Interviews

Twelve student volunteers enrolled in algebra-based physics were randomly selected from 46 volunteers. Students present at the second lecture for the algebra-based course were given the opportunity to volunteer for this study. Students were given monetary incentive for their participation. Two groups of six students were formed based upon student schedules. These 12 participants met in their respective groups a total of nine times during the semester. One of the 12 volunteers who was selected dropped the class prior to the completion of the study. The two groups of six and five students met each week for one hour.
Table 4.1 Demographics for phase 1 participants.

<table>
<thead>
<tr>
<th>Code ID</th>
<th>Trt/ Ctrl</th>
<th>Major</th>
<th>Previous Classes</th>
<th>Physics</th>
<th>M/ F</th>
<th>Semestrer</th>
<th>Class yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>Ctrl</td>
<td>Pre-Vet</td>
<td>AP Physics</td>
<td>F</td>
<td>Fall07</td>
<td>2 yrs</td>
<td></td>
</tr>
<tr>
<td>BM</td>
<td>Ctrl</td>
<td>Athletic Training</td>
<td>None</td>
<td>F</td>
<td>Fall07</td>
<td>4 yrs</td>
<td></td>
</tr>
<tr>
<td>CB</td>
<td>Ctrl</td>
<td>Secondary Education</td>
<td>High Physics</td>
<td>M</td>
<td>Fall07</td>
<td>5 yrs</td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>Ctrl</td>
<td>Bakery Science</td>
<td>High Physics</td>
<td>F</td>
<td>Fall07</td>
<td>4 yrs</td>
<td></td>
</tr>
<tr>
<td>KH</td>
<td>Ctrl</td>
<td>Animal Science/Pre-Vetinarian</td>
<td>High Physics</td>
<td>F</td>
<td>Fall07</td>
<td>2 yrs</td>
<td></td>
</tr>
<tr>
<td>KN</td>
<td>Ctrl</td>
<td>Biology</td>
<td>High Physics</td>
<td>M</td>
<td>Fall07</td>
<td>3 yrs</td>
<td></td>
</tr>
<tr>
<td>KP</td>
<td>Ctrl</td>
<td>BioChemistry</td>
<td>Honors Physics</td>
<td>F</td>
<td>Fall07</td>
<td>2 yrs</td>
<td></td>
</tr>
<tr>
<td>LS</td>
<td>Ctrl</td>
<td>Bakery Science</td>
<td>High Physics</td>
<td>F</td>
<td>Fall07</td>
<td>2 yrs</td>
<td></td>
</tr>
<tr>
<td>G1S1</td>
<td>Trt</td>
<td>Biochemistry</td>
<td>AP Physics</td>
<td>F</td>
<td>Fall07</td>
<td>3 yrs</td>
<td></td>
</tr>
<tr>
<td>G1S2</td>
<td>Trt</td>
<td>Kinesiology</td>
<td>Community College Physics</td>
<td>M</td>
<td>Fall07</td>
<td>5 yrs</td>
<td></td>
</tr>
<tr>
<td>G1S3</td>
<td>Trt</td>
<td>Pre-Vet</td>
<td>High Physics</td>
<td>F</td>
<td>Fall07</td>
<td>3 yrs</td>
<td></td>
</tr>
<tr>
<td>G1S4</td>
<td>Trt</td>
<td>Kinesiology</td>
<td>High Physics</td>
<td>M</td>
<td>Fall07</td>
<td>5 yrs</td>
<td></td>
</tr>
<tr>
<td>G1S5</td>
<td>Trt</td>
<td>Microbiology</td>
<td>None</td>
<td>F</td>
<td>Fall07</td>
<td>3 yrs</td>
<td></td>
</tr>
<tr>
<td>G1S6</td>
<td>Trt</td>
<td>Kinesiology</td>
<td>High Physics</td>
<td>F</td>
<td>Fall07</td>
<td>2 yrs</td>
<td></td>
</tr>
<tr>
<td>G2S1</td>
<td>Trt</td>
<td>Life Science</td>
<td>High Physics</td>
<td>F</td>
<td>Fall07</td>
<td>2 yrs</td>
<td></td>
</tr>
<tr>
<td>G2S2</td>
<td>Trt</td>
<td>Biochemistry</td>
<td>2 yrs (international)</td>
<td>m</td>
<td>Fall07</td>
<td>2 yrs</td>
<td></td>
</tr>
<tr>
<td>G2S3</td>
<td>Trt</td>
<td>Kinesiology</td>
<td>High Physics</td>
<td>F</td>
<td>Fall07</td>
<td>3 yrs</td>
<td></td>
</tr>
<tr>
<td>G2S4</td>
<td>Trt</td>
<td>Kinesiology</td>
<td>None</td>
<td>M</td>
<td>Fall07</td>
<td>2 yrs</td>
<td></td>
</tr>
<tr>
<td>G2S5</td>
<td>Trt</td>
<td>Pre-Vet</td>
<td>None</td>
<td>M</td>
<td>Fall07</td>
<td>4 yrs</td>
<td></td>
</tr>
</tbody>
</table>

For the first two weeks of the semester, one of two moderators handed out a set of four similar deep-structure problems for students to work on briefly. The selected problems were often modified variations of problems asked in *Physics: Principles with*...
Applications, Giancoli, 6th Edition. All four problems for each week covered the same basic physical concept studied recently in the course, but had small differences in similarities. See Figure 4.1 for the range of topics covered in each focus group learning interview. Problems categorized as surface feature similar used contexts, or scenarios, that were identical or very similar, e.g. a box moving down an incline as compared with a block or box moving down an incline. Problems categorized as deep-structurally similar involved the same concepts and physical principles. Finally, a problem categorized as different in terms of complexity from another problem will require an additional concept or principle to solve. Distinct differences between problems in the sets presented for these focus group learning interviews were surface feature and/or complexity dependent. The sequence at which these four problems of varying differences were delivered remained the same throughout the treatment. These types of problem sequences, and their categorization names were based upon work done by Nokes and Ross (Nokes and Ross, 2007). Figure 4.2 below contains examples of problems given to students during the focus group learning interviews. A description of how the problem sequence was assembled from week to week is given below.

**Figure 4.2 Top left: Problem 1 from week 6. Top right: Problem 2 from week 6. Bottom left: Problem 3 from week 6. Bottom right: Problem 4 from week 6.**


<table>
<thead>
<tr>
<th>Problem 1</th>
<th>Problem 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A sled is initially given a shove up a frictionless 30.0 degree incline. It reaches a maximum vertical height 1.35 m higher than where it started. What was its initial speed?</td>
<td>A sled is initially given a shove up a frictionless 30.0 degree incline. It has an initial speed of 6.0 m/s. What will be the maximum change in vertical height acquired by the sled?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Problem 3</th>
<th>Problem 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>A medieval archer fires an arrow at an upward angle of 80 degrees from the bottom of a 265 meter wall. Assuming that the archer barely clears the top of the wall, what would be the required initial arrow speed?</td>
<td>A vertical spring whose spring stiffness constant is 950 n/m, is attached to a table and is compressed down 0.20 m. To what height above its original position (spring compressed) will a 3.0 kg ball fly?</td>
</tr>
</tbody>
</table>
In the sequence of problems provided to the students, problem 1 and problem 2 were always similar in terms of surface and deep-structure, but the quantity that students were asked to solve for in problem 2 replaced a quantity previously given in the problem statement in problem 1. These types of pairs are called near-miss (NM) pairs. See Figure 4.3 for a visual representation of the problem sequence description. Problem 3 remained deep-structure similar to problem 1 and problem 2, but the surface features were different. Problems pairings 1 and 3, and 2 and 3, are referred to as structure similar (SS) pairs because they share the same principles. Problem 4 had additional complexity by requiring students to utilize previously studied principles in addition to the primary concept and principle. Problem 4 is considered surface feature and deep-structure different from all other problems in the set because it varies both its scenario and requires application of additional principles for a solution. These problems are referred to as Principle Different (PD).

**Figure 4.3 Visual description of the variance of problem types given during the focus group learning interviews.**

In the third week, students were introduced to structure maps for a given section of the textbook. Like our pilot phase, the maps were pre-created by experts with the purpose of focusing training on map utility rather than map creation. See Figure 4.4 for examples of structure maps used in the first half of the semester. The nodes contained quantities and were connected to each other in one of two ways: by their sidewalls or by arrows. By connecting quantities by their sidewalls, we aimed to represent a specific association between quantities that may be written as an equation. By connecting the quantities with arrows, we aimed to represent a more general association between quantities. The arrows were used in cases where the equation could differ depending upon the problem context.
Both treatment groups were given the same instructional PowerPoint slideshow. The training slideshow explicated the creation of the map. Each equation presented in the chapter section was transformed into a visual representation of the associations between quantities. With every new equation, the visual representation of associations grew larger. Once the completed map was presented, and example problem was selected and presented alongside the completed map. The map would then begin to highlight information given in the problem statement, information implicitly known, and information asked for by the problem. Finally, the map would highlight quantities that must be calculated in order to complete the solution to the example problem. Once the powerpoint was shown, students were handed the same structure map and problem one from the problem set. Students were trained to use the structure map handed out by the moderator by marking ‘X’s of varying colors through quantities that are given in problem 1, the quantity that is asked for in the problem and quantities that must be calculated in order to progress from the given quantities to those asked for in the problem. For all future group learning interviews, students proceeded to use the maps while solving the
problem sets, but they were given time to use the map in their own way. Students were allowed to use the same printed map for all four problems given during the interview, but often students opted to take a new printout for each problem. Students were allowed and encouraged to help one another. Assistance was provided only when students were unable to help one another. Students were asked to react to the structure maps and discuss elements of the map they found useful. Students would sometimes be asked to present their structure maps and problem solutions to the group if they were quiet for too long.

The maps were restructured after completion of interview 6 to accommodate some of the students’ suggestions made during the first 6 focus group learning interviews and the first individual interview. Structure maps used for the final three focus group learning interviews became visual representations of equations and the relationships between quantities within equations for a given section of the textbook. (See Figure 4.3 for the map used during week 9.) At the end of the group interviews, students were asked to explain how they felt about the new maps and discuss the features they found most useful.

### 4.3 Results: Group Learning Interviews

Student performance on problem pairs and perceptions of the structure maps were noted for each week of the group interviews using original video and audio recordings. Students were given four (4) problems of varying surface feature differences and complexity. See Figure 4.3 above for a representation of the problems set sequence, or visit Appendix B-2 to see full problem sets for each week. Students were asked to discuss the problems solutions, contrast and compare problems, and compare the utility of the map for individual problems. The first three interviews used different protocols from the rest, involving slightly different problem sets and either no structure mapping or a general overview discussion of structure mapping. Each group interview described below highlights the significant changes in the student-student and student-moderator exchange about problem sets and structure maps. Performance on problem sets was measured by overall completion and discussion of difficulties among participants. To probe students’ reactions and use of the structure maps, each student was asked the same
two questions at the end of each group interview and any necessary follow-up questions such as “How did you like the map?” and “How did you use the map?” Student performance on the problem solving tasks and their perception of the maps are reported in this section. Discussion related to the comparison of structural and facial features of the problems will also be reported.

4.3.1 Problem performance and Student Perception for each FOGLI

Focus Group Learning Interview 1

Problem Performance: For Focus Group Learning Interview 1, or FOGLI 1, problem sets included 2 problems pertaining to kinematics. See Figure 4.5 for the problem sets and appendix B-2 for the protocols for each week.

Figure 4.5 FOGLI 1 problem set.

| Problem A: A helicopter is ascending vertically with a speed of 5.20 m/s. When the helicopter is at a height 125m above the Earth, a 3kg package is dropped from a window.  • How much time does it take for the package to reach the ground?  • What is the final velocity of the package just before it hits the ground?  If the mass of the package in problem were 6kg…  • Would that change the time it takes for the package to reach the ground? If so, how?  • Would that change the speed of the 3kg package just before it hits the ground? If so, how? If the helicopter in Problem A were descending vertically with a speed of 5.20 m/s…  • Would that change the time it takes for the 3kg package to reach the ground? If so, how?  • Would that change the speed of the 3kg package just before it hits the ground? If so, how? If the 3 kg package in Problem A is thrown | Problem B: A stone of mass 2 kg is thrown vertically upward with a speed of 18.0 m/s…  • How fast is it moving when it reaches a height of 11m?  • How long does it take to reach this height?  • Is there a unique answer to part (b)? Explain. If the stone thrown vertically upward in Problem B is replaced with a softball of mass 0.1 kg…  • Would that change the time (compared to problem B) it takes for the object used to reach the 11m? If so, how?  • Would that change the speed (compared to problem B) of the object as it reaches 11m? If so, how? If the stone in Problem B is thrown vertically downward with a speed of 18.0 m/s from a height of 120 m above ground  • Would that change the time (compared to problem B) it takes for the object used to reach the 11m? If so, how?  • Would that change the speed (compared to problem B) of the object as it reaches 11m? If so, how? |
Problem A:

downward from a window with a velocity of 4 m/s from the helicopter ascending at 5.20 m/s at a height of 125m....

- Would that change (compared to problem A) the time it takes for the 3kg package to reach the ground? If so, how?
- Would that change (compared to problem A) the speed of the 3kg package just before it hits the ground? If so, how?

Problem B:

<table>
<thead>
<tr>
<th>Each problem was selected from the chapter problems presented in the course textbook and modified to include multiple scenarios involving the same context. Students were given one problem at a time and asked to read the problem statement, describe the physical system in the problem, state their assumptions for solving the problem, and describe how they would solve the problem without actually solving it. Problems were discussed as a group after each student was given ample time to complete the full procedure for each problem part. Students were not given structure maps to work with for this group interview. Only three (3) of 12 students completed the procedure without solving the problem first. All three students were in group 2, and all 3 assisted one another through general discussion of the tasks. Below is an example of some student conversation regarding one of the problems given.</th>
</tr>
</thead>
<tbody>
<tr>
<td>G2S2 “Isn’t change in y 125 meters?”</td>
</tr>
<tr>
<td>G2S4 “Well its ascending, so part of it will be your y initial and part of it will be your y...like if you separated it into two different problems.”</td>
</tr>
<tr>
<td>G2S1 “Yeah”</td>
</tr>
<tr>
<td>G2S4 “Isn’t it dropped at a 125 m?”</td>
</tr>
<tr>
<td>G2S1 “Well your initial velocity is the same as the helicopters...like the example...like my recitation teacher says like if you throw something out a car window and you’re like driving really fast, it’ll still keep going forward a little bit, but it’s just not going as fast as the car. you know?”</td>
</tr>
</tbody>
</table>
G2S4 “Ok, but no air resistance, like if gravity pulled slower that normal and we could see it.”

G2S2 “Haha. so like the rock is just floating next to the car until it hits the ground?....reaally sloooowwlly…haha”

Three of the remaining nine students that completed the solution prior to answering our protocol questions were a part of group 2. Two of the three went back to describe the physical system and state their assumptions, only after being prompted to do so by the moderator. The third and final student of group 2, referred to as G2S5, was the only student to complete the solution to both problems correctly, and determined that it was unnecessary to go back and complete the original instructions because he had “already solved the problem.”

All six (6) students from group 1 attempted to solve the problem first and were unable to complete the solution to any parts of the problem. All 6 students cited problems with the problem statement, stating that the helicopter would not be able to remain stationary, or hover over a particular distance. Five of the six disagreed about the preferred coordinate axis.

G1S3 “It’s not hovering, it’s moving horizontally because it’s a helicopter.”

G1S5 “No, but…the helicopter is ascending so it has to be moving vertically.”

G1S3 “I don’t think this problem is possible. You can’t have a hovering helicopter.”

[Students shake their head in agreement with the last statement]

Overall, 10 of the 11 participants had difficulty solving the problems given. The scenario of the helicopter provided challenge with respect to the real-world nature of the context. The helicopter’s vertical movement was sufficiently different from previous one dimensional motion problems involving motorized vehicles. Students were unable to complete their depiction of how air resistance and proper helicopter motion might affect a package dropped or thrown from a window. The 11th student, a member of group 2, was capable of solving all parts of both problems, but did not communicate with others in his group. When asked to reflect on the assumptions made for each part of problem A, the student voiced his concern over being asked to include unnecessary steps in the problem solving process. The student was then asked to reflect on the assumptions he had made
such that others in his group may understand his internal processing of information, but
time ran out of the interview.

*Focus Group Learning Interview 2*

**Problem Performance:** Group Learning Interview 2, or FOGLI 2, problem sets
included nine problems, made up of three sets of three problems. Each problem set
included near-miss (Problem 1 and 2) and structurally similar pairs (problem 1 and 3).
See appendix B-2 for full protocols.

*Figure 4.6 An example of the FOGLI 2 problems sets.*

<table>
<thead>
<tr>
<th>Problem 1:</th>
<th>Problem 2:</th>
<th>Problem 3:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A car starting at a velocity of 12 m/s accelerates at a rate of 6.0 m/s² to a velocity of 24 m/s. How far did it travel?</td>
<td>A car accelerates at a constant rate from rest to a velocity of 20 m/s over a distance of 120 m. What is the acceleration of the car?</td>
<td>A helicopter hovering at rest accelerates at a constant rate of 5 m/s² in the horizontal direction. How far does it travel to reach a speed of 50 m/s?</td>
</tr>
</tbody>
</table>

The problems in the first two sets were simpler than the first interview set, but
there were more of them, making it difficult to solve all of them before answering
questions about each of their characteristics. Nonetheless, all student participants solved
the problems in the first two sets and agreed upon answers.

**Similarities and Differences:** Students were asked as a group to describe the
similarities and differences between each problem, and what changes would have to be
made to make problems appear more similar. Students in both groups completed the first
set of problem easily, and answered the questions regarding similarities and differences
as a collective in agreement. Student believed the first two problems were very similar
because they used the same equations, and different because the problems were
interchanging a car with a helicopter. Examples of student conversation are given below:

G2S5 “These two are the same, they use the same equations.”

G2S4 “This one does too. Except it’s a helicopter. You just need to make this one a car and it would be the same as the ones above.”

After completing the second set of three problems, students began to voice concerns over
the ease of the task. Again, students answered as a collective.

G2S5 “These are all the same.”
G2S1 “These are even more similar than the last ones, all your changing in the third problem is the direction. That’s it!”

One student in group 1 had trouble understanding the motion of a ball and the others in the group explained,

G1S3 “why do I…where does that [term] come from?”
G1S1 “Since your throwing it up you have to deal with gravity. So that’s your acceleration.
G1S3 “is it still in one dimension? no?”
G1S2 “its only vertical…[a long moment passes] are you still confused”
G1S3 “oh..noooo….hehe..im stupid.”

Students were dismissed from the interview following discussion of the second set of problems. The interviews were slotted to take no more than 50 minutes of the participants’ time and both groups required most of the 50 minutes to complete the first two sets. This is the only interview of the nine interviews that included significant discussion of similarities and differences between problems. The moderators failed to insist on this component of the interview protocol for the remainder of the interviews.

**Focus Group Learning Interview 3**

Students are introduced to force in one and two dimensions in FOGLI 3. Four problems are created for this interview, but the focus is on structure mapping training. See Figure 4.7 below for referencing the problem set and force structure map. The structure mapping training may be found in Appendix B-1.
Problem 1:  
A person applies a force $F = 20 \text{ N}$ to a box of mass $m=10 \text{ kg}$ sitting on a horizontal surface with a coefficient of friction between the box and the surface to be $\mu_k =.10$. Determine the acceleration of the system.

Problem 2:  
A person applies a force to a box of mass $m=10 \text{ kg}$ sitting on a horizontal surface with a coefficient of friction between the box and the surface to be $\mu_k =.10$. The box accelerates at a constant rate of $a = .5 \text{ m/s}^2$. Determine the amount of force applied by the person to the box.

Problem 3:  
A person is skydiving 10,000 ft in the air and the mass of the person and the parachute is $m=70 \text{ kg}$. Assuming that the upward force of wind resistance is $F_{fr}=650\text{ N}$, determine the acceleration of the skydiver.

Problem 4:  
A box lies on a plane tilted at an angle of $\theta = 15^\circ$ to the horizontal, with coefficient of kinetic friction between the plane and the box $\mu_k=.10$. Determine the acceleration of the box as it slides down the plane.

**Problem Performance:** The first three problems in the interview set required forces in one dimension. The fourth problem was not much more complicated than previous problems, but an incline was added, requiring students to break up the forces into components. With exception to G2S5, all students in both groups participated in working through a structure map training guide, and used the remainder of the session to practice...
on a training problem. The training guide and practice problem were implemented as viewable in appendix B-1. Students were given the set of four problems, but there was insufficient time for students to consider the similarities and differences between problems or discuss their individual solutions with one another. All students completed the first two problem solutions correctly in their notebooks.

**Structure map Perception:** The group discussion following the training focused on initial difficulties while using the map to identifying quantities used in the problem. Difficulties which arose were mostly trivial, such as whether the convention chosen in the training guide should be the same as used by each individual. The researchers felt it was necessary to allow students to create their own marking systems if they were uncomfortable with the one shown in the Powerpoint training. It was made clear during this interview that students could mark the structure map in any way they wanted just as long as they made a key describing the types of markings for the moderators. The other concerned voiced by students was related to whether quantities known to cancel should be marked or not marked on the structure map. The moderator instructed students to mark as they felt best for their own map use.

G1S2 “The mass doesn’t seem to matter because they’re going to cancel out, from the x and y. I tried to show that on the map, showing both. I didn’t really know what to do with the angle, I just crossed out one because I knew I had an angle.”

**Focus Group Learning Interview 4**

Students continued to work with forces in week 4, but they had moved to circular motion in class. The pre-created map as shown in Figure 4.8, like the previous map shown in Figure 4.7, illustrates both coordinates using map symmetry. It also includes centripetal acceleration components. Figure 4.8 also includes the problem set used for the fourth focus group learning interview.
Problem 1:
A person is placed in a “rotor-ride” at a carnival. The room has a radius of 5 meters, and a rotation frequency of 0.5 revolutions per second when the floor drops away. What is the minimum coefficient of static friction required so that the person does not slip down?

Problem 2:
A person is placed in a “rotor-ride” with a radius of 5 meters, and the coefficient of static friction is $\mu_s = 0.1$, what is the rotation frequency in revolutions per second required so that the person does not slip down?

Problem 3:
A coin is placed 10 cm from the rotating axis of a turntable of variable speed. When the speed of the turntable is slowly increased, the coin remains fixed on the surface until a rate of 0.5 revolutions per second was reached. What is the coefficient of static friction between the coin and the turntable?

Problem 4:
A car travels around a curve of radius 50 meters with a speed of 16 m/s. If the curve is banked on an angle of $\theta = 20^\circ$ from the horizontal, what is the coefficient of static friction required to keep the car on the track?
**Problem Performance:** All but two students of group 2 were able to complete solutions to two of the four problems. Students identified quantities given and asked for on the map, but determined that it would be easier to solve the problem first and then go back to find the quantities and associations between quantities necessary for completing a solution. This decision was made prior to completing any of the problems in the set, and the valiant effort to use the maps prior to solving the problem lasted approximately five minutes before a general declaration of dislike was made from the group.

G2S1 “I’m not doing this part. I don’t think that way, I need to solve it out first.”
Moderator “Ok”
G2S1 “I can put the quantities asked for down first and that might help find the formula needed. maybe”

It was determined from the conversation which followed that students were having difficulty with the concept of a centripetal force. In group 1, students were reintroduced to the concept of centripetal acceleration, and problem 1 was used as an example of how to discern from the many forces present in the given scenario. In group 2, students G2S4 and G2S5 were asked to go over their solution with their group mates, discuss how they interpreted the scenario, and what kinds of forces must be acting in the given situation.

One student in group 1 became distressed even after the explanation of the problem solution by the moderator. It was apparent that this student, among other in these groups, had not understood a centripetal force would need to exist in real life. It is also interesting to note that the discussion about centripetal forces was not required for most students to solve the first two problems. It was only after students were asked to explain their solutions to one another that a meaningful conversation was prompted:

Moderator “What kinds of example did [instructor name] use to demonstrate centripetal force?”
G1S2 “He swung a bucket of water over his head.”
Moderator “Ok. How did you know there must be a centripetal force involved?”
[Lots of blank looks.]
G1S2 “Don’t know. Because it’s swung really fast.”
Moderator “How about if I took a bucket of water and tipped it off the table really fast and then tipped it back up again? Would the water remain inside?”

G1S1 “No, you need to have some r”

G1S3 “Oh yeah, well maybe you could do it if you did it really really fast. haha nevermind, you still don’t have an r, so no.”

Moderator “What is it that makes the distance r so important?”

G1S3 “It’s in the equation.”

G1S1 “If it’s a force that only occurs in circular motion, then you need the thing to move in a circle, and that circle needs some r.”

G1S2 “Wait, so all spherical movement or whatever has an extra force that other movement does not?”

G1S3 “No”

G1S2 “It would have to be.”

The group interview time ran out prior to discussing any other problems students may have completed, but overall, the two groups did not complete more than two problems on average.

**Structure map Perception:** Only two students, one in group 1 and the other in group 2, stated they found the map useful. These students believed the map organized information from the problem statement so that connections between quantities could be readily seen.

G1S2: “I do like the concept map because there are some equations involved in physics that it’s just like, when you need that and how do you use it and why and kind of I don’t know, just organizes it better so that you can look at what connects with what.

Most students, nine of the 11 total, voiced a concern that the map was hard to follow. Some were more specific in stating that not all of the information needed could be displayed on the map while others agreed. One student (G2S3) stated:

“I don’t know how to mark it on here exactly. Like, I couldn’t figure out how to relate like the terms on here to the problem.”
Two students, one from each group, also felt that the map was difficult to use without an equation sheet nearby. Most other students within both groups agreed with this claim as well. Students were given the option to modify the map on the spot, but often felt no desire to salvage any part of the map. The students used the map only after solving the problems when they were reminded to do so by the facilitators. One of the five students in the second group, G2S5, refused to use the map at all because he felt the map served no purpose other than to confuse him.

**Focus Group Learning Interview 5 & 6**

During group learning interview 5 and 6, the same structure map was presented on work and energy. The map’s physical layout was similar to maps previously used, but because this map covered work and energy, there was a temporal symmetry to the map that was not previously available for other physics principles like kinematics or dynamics. The left side of the map contained all initial quantities and the right side contained all final quantities. Figure 4.9 includes the work energy structure map and the problem set for FOGLI 5. Figure 4.10 includes the problem set for FOGLI 6.
Figure 4.9 Work and Energy structure map and problem set for FOGLI 5.

<table>
<thead>
<tr>
<th>Problem 1:</th>
<th>Problem 2:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 0.10 kg arrow is fired from a bow whose string exerts an average force of 110 N on the arrow over a distance of 0.8 m. What is the speed of the arrow as it leaves the bow?</td>
<td>A 0.10 kg arrow is fired from a bow with a speed of 50 m/s over a distance of 0.8 m. What is the average force exerted on the arrow by the bowstring?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Problem 3:</th>
<th>Problem 4:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 0.14 kg baseball exerts an average force of 300 N on a fielder’s glove, moving the glove backward 0.25 m when the ball is caught. What is the speed of the ball?</td>
<td>A 1200-kg car rolling on a horizontal surface has speed of 18 m/s when it strikes a horizontal coiled spring and is brought to rest in a distance of 2.2 m. What is the stiffness constant of the spring?</td>
</tr>
</tbody>
</table>

Figure 4.10 FOGLI 6 problem set.

<table>
<thead>
<tr>
<th>Problem 1:</th>
<th>Problem 2:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A sled is initially given a shove up a frictionless 30.0 degree incline. It reaches a maximum vertical height 1.35 m higher than where it started. What was its initial speed?</td>
<td>A sled is initially given a shove up a frictionless 30.0 degree incline. It has an initial speed of 6.0 m/s. What will be the maximum change in vertical height acquired by the sled?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Problem 3:</th>
<th>Problem 4:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A medieval archer fires an arrow at an upward angle of 80 degrees from the bottom of a 265 meter wall. Assuming that the archer barely clears the top of the wall, what would be the required initial arrow speed?</td>
<td>A vertical spring whose spring stiffness constant is 950 N/m, is attached to a table and is compressed down 0.20 m. To what height above its original position (spring compressed) will a 3.0 kg ball fly?</td>
</tr>
</tbody>
</table>
**Problem Performance:** Work and Energy problem sets were used for group learning interviews 5 and 6. FOGLI 5 made use of work-energy theorem problems and FOGLI 6 employed conservation of energy problems. There is a noticeable difference in problem solving performance after FOGLI 4. All students completed all four problems in both FOGLI 5 and FOGLI 6 with exception to problem 3 in FOGLI 6. The problem was not solvable using work and energy principles, only kinematics. One student from each group was able to determine that problem 3 was not solvable using recent work and energy methods. The student participating in group two further pointed out that the assumption regarding the archer barely clearing the top of the wall would make the final velocity in the y-direction zero, but not in the x-direction. This student went on to describe how kinematics might be the best method for solving this problem.

For the remainder of the problems, all students were capable of explaining their solutions correctly, discussing their map usage, and briefly discussing problem similarities. The structure map was heavily used by students to explain solutions.

G2S3: “The map makes this so much easier to explain. We are given these quantities I’ve marked in blue, and we want to find this quantity in red. In order to get from blue to red, I need to calculate the total energy in the sled before it goes up. It is the same as the total energy at the up position, but its not moving anymore so its all this guy [gravitational potential] right here!”

One difficulty arose during the discussion of problem 1 in FOGLI 5 stemming from an initial confusion between work-energy theorem and conservation of energy, but was worked out through discussion with peers. Students completed one homework and lecture on the work-energy theorem, and a lecture on conservation of energy prior to the FOGLI 5. Student G2S3 wanted to treat a bowstring in problem 1 like an elastic band, but was not given the right information to simply use conservation of energy. When told by a classmate that they must apply the work-energy theorem, the student was unsure of how the work-energy theorem took the initial energy into account without having mathematical expressions in terms of potential. The participant G2S3 then determined that since the potential energy would be exactly equal to the kinetic energy after release,
they could plug in the given velocity as the final velocity. It took little prompting from another student in the group to get G2S3 back on track.

G2S3 “But don’t you have an initial velocity?”

G2S5 “Your initial velocity is zero. you have some potential, but its not moving”

G2S3 “Oh…oh!”

**Problem Similarities:** During the brief similarity comparison in FOGLI 5, students would point to the quantities that were given and asked for, and note that they were switched between problems with a hand gesture. By problem 3 in FOGLI 5, students made the decision to highlight the similarities between problems using their own map language. One student would draw a double arrow between quantities that were flipped, and double circle nodes that represented a new concept introduced in problem 3. Another student would color nodes in if they were the variable which changed or was added in the problem.

**Figure 4.11 Example of students’ map markings (darker color represents the difference in quantities between two maps).**
**Student Perception:** There was a significant difference in student feedback on this map compared to the previous maps. All 11 students favored the work energy map over the previous kinematics and force maps, but only 10 of 11 used the map during the interviews. When we asked each student individually about their thoughts on the structure map, the one student that did not use the map stated, “I did not need it. Why would I use it?” For future reference, I will refer to this student as G2S5.

Other participants were also asked to explain why they liked the work energy map over the previous map. A typical response is below:

G1S1 “I feel like this one (work-energy map) you’re just looking for your potential energy final, like I feel like it’s just easier to focus in on that area [of the map] and how you would lead there, other than the other one (force map) that’s just like you have [a quantity] down here but you feel like you have to go through all the other bubbles.”

G2S4 “Like this (force map) it’s all one big thing, but for this (work energy map) you can follow along so you can go from this to get this and …like you can follow the arrows on this one.”

All other participants, with the exception of G2S5, used the map and found it easier to navigate between quantities that were given and those that were asked for in the problem. Some referred to the work energy map as being similar to a “road map.” They were capable of selecting all values that were given in the problem by circling them and selecting the value that was asked for by crossing an ‘X’ through the value. They then established a clear path following arrows which led from the quantities given in the problem to those asked for in the problem. Though this map was better received than previous maps, students still wanted equations to be provided. Two students also stated that they would prefer to see units included with the maps.

G1S3 “…units, like what the units should be, like knowing what each value should have for units, so I know when I do the problem I’m not missing that.”
Focus Group Learning Interview 7 & 8

Students were given some time away from the focus group learning interviews such that they could complete the first individual interview. The researchers took this opportunity to filter through the comments made by students about the structure maps. Students voiced opinion over several map components, most often during force and kinematics maps. Students believed that it was difficult to progress from one point in the map to another smoothly, as the associations between quantities, though obviously shown on the map, were unclear. It was also determined that students, while paid to participate in these focus groups, were unwilling to adapt pre-created maps to fit their own needs. Students also stated that even the most useful work energy map could use equations worked into the visual representation. Researchers decided an alternative to the current structure maps was to include the equations in the nodes, instead of the quantities. The arrows would no longer represent associations between quantities, but associations between equations. A new map using this design was created for focus group learning interview 7.
Problem 1:
A pumpkin with mass 3.0 kg is attached to a spring with stiffness constant $k = 280$ N/m and is executing simple harmonic motion. When the pumpkin is 0.020 m from its equilibrium position, it is moving with a speed of 0.55 m/s. Calculate the maximum velocity attained by the pumpkin.

Problem 2:
A pumpkin with mass 3.0 kg is attached to spring is executing simple harmonic motion. When the pumpkin is 0.020 m from its equilibrium position, it is moving with a speed of 0.55 m/s. The amplitude of motion is 0.060 m. Calculate the correct spring stiffness constant.

Problem 3:
A spring of a toy popgun is compressed 0.200 m to “load” a 0.180 kg ball. Assuming the spring has a stiffness constant $k = 110$ N/m and leaves the gun with a speed of 0.25 m/s, what is the maximum velocity attained by the ball?

Problem 4:
A 0.60 kg mass vibrates according to the equation $x = 0.45 \cos(6.40t)$, where $x$ is in meters and $t$ is in seconds. Determine (a.) amplitude, (b.) frequency, (c.) total energy, (d.) how far away from equilibrium the ball is after 0.20 s, and (e.) the velocity of the mass after 0.20 s.
<table>
<thead>
<tr>
<th>Problem 1:</th>
<th>Problem 2:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 0.755 kg mass at rest on the end of a horizontal spring (k=124 N/m) is struck by a hammer, so that the maximum displacement from the rest position is 0.23m. Determine the..</td>
<td>At t=0, a 0.755 kg mass at rest on the end of a horizontal spring is struck by a hammer, so that the maximum displacement from the rest position is 0.23m. The maximum acceleration of the mass is 38 m/s², determine the stiffness constant of the horizontal spring.</td>
</tr>
<tr>
<td>a) maximum velocity</td>
<td>a) maximum velocity</td>
</tr>
<tr>
<td>b) maximum acceleration.</td>
<td>b) maximum acceleration.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Problem 3:</th>
<th>Problem 4:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 0.30-kg bullet is loaded into a spring gun, and rests on the end of a horizontal spring (k=1.2x10⁶N/m). When the hammer in the revolver comes down, a force strikes the spring displaces the bullet a distance of 0.06m from the rest position. Determine the..</td>
<td>A pianist is tuning the “middle C” on his piano. As he pushes down on the key, the piano cord vibrates in simple harmonic motion at a frequency of 268 Hz. At t=0, the amplitude is A=1.5x10⁻⁴m.</td>
</tr>
<tr>
<td>a) maximum velocity</td>
<td>(a) What is the velocity as a function of time?,</td>
</tr>
<tr>
<td>b) maximum acceleration.</td>
<td>(b) What is the acceleration as a function of time?</td>
</tr>
</tbody>
</table>

**Problem Performance:** Group learning interviews 7 and 8 covered simple harmonic motion. Ten of eleven students in the two groups completed all four problems. Student G1S2 was unable to complete problem 4 in interview 7 because he continually remained at least one problem part behind others in the group. Once the group completed two parts of five in problem 4, G1S2 asked for assistance, and the group collectively organized a brief synopsis of the problem solution to that point. Student G1S2 worked by himself in an attempt to learn from the first two parts of the problem, but did not have enough time to complete the solution in the interview. Student G1S2 stated he would continue to work out the problem later in the day for his own benefit. The rest of the group continued to work together, and made several comments that implied mechanical plug and chug solution procedures:

- G1S3 “Which one is omega?”
- G1S1 “Yeah, it’s this guy, just match ‘em up”
- G1S4 “Uhh..yep. And that’s the amplitude...easy to identify our equation”

[work together on solving problem 4]

Student G1S2 remained consistently behind in group learning interview 8, but was capable of completing the problems with minor assistance from the group. During
interview 8, he noted that he was behind in his coursework due to a family concern. The moderators did not step in for anything more than minor clarification for both of these interviews.

**Problem Similarities:** Unlike the previous two interviews, students were no longer using the map to identify similarities and differences between problems. Students were also not prompted by the moderators to discuss similarities and differences between the problems since much of the time was spent talking about the new map design.

**Perception:** The map contained equations in the nodes, while the arrows represented relationships between quantities within the equations. Initially, the map was viewed as too complicated by students in both groups.

G2S3 “Well I felt like I needed it in problem two…I don’t know, it’s just a lot of arrows…. a lot more stuff I guess. It is intimidating.”

Only two students in group 1 and one student in group 2 initially used the arrows between quantities to guide them to a solution. Their solution was correct, but not the most efficient way to solve the first problem. The student in group 1 recognized that the solution was unnecessarily long, but felt like she learned how the quantities related better.

As the session progressed to problem 2, all 11 students stated that they liked the map design, while 10 students found the new map to be useful while solving problems. Students liked having the equations given directly on the map. Many (7 of 11) stated that the arrows connecting quantities across equations were very helpful.

G1S3: “[This map is] a lot easier to use. I don’t have to like look up a bunch of different equations like, oh I don’t have that… you can just see how everything relates and what you have and how it works together.”

Similarly, G2S3 no longer felt the map was intimidating, determining that no arrows between quantities using similar notation (i.e, \(x\) verus \(x_{\text{max}}\)) was a good indication that those were not identical values. G2S5 also decided (for the first time) that he liked the map and used it for problem 4, but generally did not prefer using any map. Here is a small segment of the group 2 interview:
Moderator  “What did you think of the map after problem 4?”
G2S2  “I like it.”
G2S5  “Yeah.”
G2S3  “Used it a lot. It’s nice.”
G2S2  “I might actually be putting it on my cheat sheet for the test…It’s easy to understand.”
Moderator  “Okay. Were the arrows helpful?”
G2S3  “Yeah, because if you didn’t know what you were doing to an extent, but you know kinda what you’re doing, you could be like this problem it (v) doesn’t link to this (v_max) because your arrow isn’t there. I kinda looked at it that way.”
G2S5  “This is good, but can I say my personal opinion?”
Moderator  “Of course.”
G2S5  “I prefer to work without maps. If you know the equation, you know the variables, then there’s no need to see this thing [structure map], like that’s my…I don’t know.”

During this final group interview, three students, one from group 1 and two from group 2, made it clear that the new map would be added to their ‘cheat sheet’ for their final examination along with the work-energy map. The ‘cheat sheet’ could be any 8 ½ x 11 sheet of white paper with notes or equations written on both sides. Problem examples were not permitted. Multiple students asked for extra copies of the work and energy map while also taking extra copies of the vibration and waves map.

After this interview, moderators investigated how students’ perceptions of the maps aligned with their overall class performance. Students that believed that the maps were useful and also determined that they could be used on their cheat sheets, were average students with grade distributions in the low B or mid to high C range. Students that were performing below average in the class and also participating in the FOGLI’s did not perceive the maps as useful, but did believe the maps could be useful. Students participating in the FOGLI’s that were also ‘A average’ students also stated that the maps were not useful to them, but again, that
the maps could be useful to others. Student G2S5 was an example of an ‘A’ student. This pattern was determined from the FOGLI participants average scores in the class just prior to the final examination.

**Focus Group Learning Interview 9**

Students covered fundamental frequencies and standing waves prior to the final focus group learning interview. Another map was created for this material using the new design, though there were only four significant equations necessary for this section. The map and problem sets are shown below. A fifth problem is also created for this interview which aimed to incorporate additional concepts previously studied and present a problem of higher complexity, but students were never given the fifth problem due to time constraints and lack of motivation on the final date. The fundamental frequency structure map and problem set may be seen below. The fifth problem is not shown in figure 4.13, but is apart of the full protocol presented in Appendix B-2.

**Figure 4.14 Fundamental frequency structure map and problem set for FOGLI 9.**

<table>
<thead>
<tr>
<th>Problem 1:</th>
<th>Problem 2:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A guitar string is 90 cm long and has a mass of 0.0036 kg. The distance from the bridge to the support post is L=62 cm, and the string is under a tension of 520 N. What are the frequencies of the fundamental and first two overtones?</td>
<td>A 90 cm long guitar string vibrates at 300 Hz as its fundamental frequency, and is under a tension of 530 N. The distance from the bridge to the support post is L=62 cm. What is the mass of the guitar string?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Problem 3:</th>
<th>Problem 4:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A woodpecker lands on a power line strung between two poles and begins pecking at the line at his feet, forcing the power line to vibrate. The powerline is 10</td>
<td>One end of a horizontal string is attached to a small amplitude mechanical 60 Hz vibrator. The string is 2 meters long and has a mass of ma= 0.0008 kg. The string</td>
</tr>
</tbody>
</table>

---

86
meters long and has a mass of 3 kg. The distance between the woodpecker the pole is 8 meters. The power line is under a tension of 150 N. What are the frequencies of the fundamental and first two overtones? 

passes over a pulley, a distance L=1.50 m away, and weights are hung from this end. (see figure below) What mass, m, must be hung from this end of the string to produce…

(a) One loop of standing wave?
(b) Five loops of standing wave?

[HINT: Assume the string at the vibrator is a node.]

Problem Performance: Group learning interview 9 covered waves and fundamental frequencies. Students consistently work through solutions to the problems using the structure maps, but again there was no discussion of problem similarities. There was an obvious lack of motivation in the last interview for group 1. Students completed all of the 4 problem solutions and used the maps to their own advantage, but relied on each other to confirm answers or procedures to the solutions rather than confirm their solutions with conceptual reasoning.

G1S3 “Which equation are we using?”
G1S4 “This one”
G1S1 “Oh”
G1S2 “We did one like this in class”
G1S3 “So in the second overtone the wavelength should just be L?”
G1S4 “Well I think you can just take it like 2F. just take the frequency times two. And the next one you take the frequency times three.”

After completing the first two problems, moderators instruct students in group 1 to explain why two separate solution methods were used and are both correct. After a moment, G1S3 begins to respond:

G1S3 “It would work the same way if you took the velocity …[mumbles] and divide by L.”
Group 2 students were also less motivated in the final group interview. Both groups went on conversations related to the upcoming examination. The fifth problem was put on the table five minutes before the end of the hour, but students remained intent on tangent conversations. The moderators were unable to motivate a change in conversation. No one from group 1 looked at the fifth problem, and only G2S5 from group 2 looked at problem 5. He began to work the problem, only to be stopped by G2S4 asking a question related to the upcoming in-class examination.

**Perception:** The only conversations which remained focused were students’ comments related to the perceived usefulness of the new structure map. Student G2S5 was an unexpected catalyst in discussion of structure maps, referring to the new map as “strange” and useful.

Moderator “Why would you describe this map as strange, [student name]?”
G2S5 “It’s not like anything I’ve seen or used before, so it’s strange. I prefer this map type to the previous ones, before our independent interviews. They are an equation sheet, but with additional relationships also.”
G2S4 “Yeah, this is totally on my final exam cheat sheet too.”
Moderator “Why not on your cheat sheet for your exam next week?”
G2S2 “[Instructor name] said we weren’t covering this section for this exam.”

### 4.3.2 Summary of FOGLI Results

Student participants appeared to be well-motivated and consistently well-focused on problem tasks and discussion over eight of the nine Focus Group Learning Interviews. The focus group learning interviews were used to create an environment where students could collaboratively work through problems of varying similarities using a structure map as a facilitator for explicit organization and processing of problem information. For each problem, students were asked to highlight on a pre-existing structure map: information given in the problem statement, information implicitly known, and information asked for by the problem. Students were also asked to highlight quantities that must be calculated in order to complete the solution. For the remainder of each
interview, problems were solved, problems were compared and contrasted, and the utility of the structure map was discussed.

From the data collected from our video and audio of each individual interview, we were able to address each of our original research questions. Each of these research questions will be answered here.

**How do students perform on the problems sets given in the focus group learning interviews?**

As the semester progressed past the first four weeks, students were capable of completing the problem solutions and did a fair job of assisting peers in need of guidance. Difficulties arising from the concept of centripetal motion during week 4 may have been at least partly due to incorporating problems before students were given the chance to complete their class homework assignment. All other concepts covered in the FOGLI’s were covered in homework prior to the group meeting. Anecdotal information collected from the teaching assistants suggest that the students also did not cover the concept completely in lecture.

**How do students compare and contrast problems of varying similarity?**

Students primarily cited the use of identical equations as being a prominent similarity between two problems. Students cited surface features, such as a helicopter or a car, as the differences between problems. Students never described the similarities and differences between problems 3 and 4 of any given set. Similarities and differences described by students were never distinguished as ‘important’ or ‘unimportant.’ It might be appropriate for future studies to ask students to clarify how significant a similarity or difference is to the solutions of the problems.

**How useful were the individual maps perceived while problem solving?**

Force structure maps were disliked by the majority of students in the focus group learning interviews. Work and energy maps were perceived as more useful as compared to the force maps because the progression from the quantities given to the quantities asked for could be seen directly on the map. The force map did not seem to have a clear connection between the quantities given and quantities asked for. The redesigned maps
covering vibrations, waves, and fundamental frequencies were also well received. Students determined that the redesigned maps were useful because they acted like equation sheets, but with additional information regarding the associations between quantities within individual equations. It is important to note that though the majority of students perceived the maps as useful, only three of the 11 students used the maps on their own cheat sheets, and only C and low B average students expressed interest in using the maps for themselves.

**How do structure maps affect collaborative problem solving?**

It is difficult to compare student performance on problem set solutions to students use of the structure maps, as there is no clear evidence of whether the difficulties corresponding to problem sets are due to difficulties corresponding to the given structure map, or vice versa. It is also unknown as to whether students may be predisposed to certain preferences such as temporal symmetry versus spatial symmetry in problems or maps. Structure maps that are declared by the majority as being useful and problem sets that are correctly and fully completed do in fact coincide for FOGLI’s 5 through 9. Unfortunately there are too many dependent variables interchanging from week to week to account for this coincidence.

### 4.4 Methodology – Individual Interviews

Each participant was asked to participate in two individual interviews over the same semester. They met once with one of the moderators at the mid-point of the semester, and again, at the end-point of the semester. The protocols for both interviews contained three non-traditional problem solving tasks. The non-traditional tasks included problem posing (Mestre, 2002), text editing (Low & Over, 1990), and physics Jeopardy (Van Heuvelen, 1998). Problem posing requires students to take a given scenario and pose a problem dependent on that scenario. Often the problem scenario may include more than one specified principle. This task was designed by Alan Van Heuvelen as a measure for probing students understanding of physics concepts and their ability to transfer their knowledge to novel contexts. Text editing requires students to determine whether a given problem statement is missing information, contains irrelevant information, or contains
sufficient information necessary to solve the problem. Text editing is defined as a 
measure of schematic knowledge, or students’ mental organization of knowledge 
elements and associations between such elements. Finally, Physics Jeopardy requires 
students to identify a scenario that best matches a given complete or partial solution. 
Physics jeopardy is a measure of students’ understanding of the mathematical 
representation of a given physical process.

Eight students of equal to similar grade distribution that volunteered in the 
beginning of the semester were called back to participate in the individual interviews. 
These 8 students served as a baseline, or a group we could use to compare with our own 
cohort with respect to performance on the individual interviews. Baseline students were 
required to do the same tasks as our treatment participants.
4.4.1 Mid-semester individual interview

Interviews conducted at the mid-point of the semester included four problem tasks, three include the non-traditional tasks described earlier. These tasks could not be completed using any mechanical solving routine. Both interview materials may be found in appendix B-3.

Task 1 – Problem Posing

Problem Posing was assigned as task 1 of the mid-semester interview. Students were given a picture which might depict a possible physical scenario, and asked to create their own physics problem, or problems, based upon the situation. They were told explicitly that they could use anything they have learned in physics up to that point in the semester.

Figure 4.15 Problem posing task assigned for individual interview 1.

Take a look at the picture below. Create your own physics problem based upon this situation. You may use anything that you have learned from General Physics.
Task 2 – text editing

The Text Editing task was assigned as task 2 (Low and Over, 1990). Given a problem statement, students were asked to find irrelevant or missing information, or if none existed, declare the problem sufficient.

Figure 4.16 Text editing task assigned for individual interview 1.

Please specify whether the problem statement above provides sufficient, missing, or irrelevant information for applying toward a solution.

A 72 kg motorcycle daredevil is attempting to jump across as many buses as possible. The takeoff ramp makes an angle of 18° to the horizontal, and the landing ramp is identical to the takeoff ramp. The buses are parked side by side, and each bus is 3.5 meters wide. The cyclist leaves the ramp with a speed of 30 m/s. What is the maximum number of buses over which the cyclist can jump?

Task 3 – Physics Jeopardy

Task 3 was a physics jeopardy task. Students are given a solution or partial solution to a problem, and asked to identify a scenario that corresponds to solution.

Figure 4.17 Physics Jeopardy task assigned for individual interview 1.

The given information below is a worked-out solution to a Physics Problem. First identify what concepts are used to solve the problem. Then describe a real-life situation that best fits this solution.

\[
\frac{mv_1^2}{2} + mgy_1 = \frac{mv_2^2}{2} + mgy_2 \rightarrow \frac{mv_1^2}{2} + 0 = 0 + mgy_2
\]

\[
\frac{(55 \text{ kg})(60 \text{ m/s})^2}{2} + 0 = 0 + (55 \text{ kg})(9.8 \text{ m/s}^2)y_2
\]

Task 4 – Creation of a concept map

The fourth task was to create a concept map, not a structure map. Students were given an example from biology, where the nodes are either types of living things or actions that describe a living thing. The lines represent the type of association between the node, e.g. a ‘living thing’ node may be connected to a ‘consume energy’ node by a
line ‘does’ line, so we may interpret this connection as a living thing does consume energy.

Students were asked to create their own concept map from the previous physics Jeopardy task. The purpose of this task was to determine whether students from our focus group learning interviews were capable of producing a concept map and whether their map might be unique as compared with the work and energy structure map presented in our most recent focus group learning interview.

**Figure 4.18 Concept map task assigned for individual interview 1.**

When you’ve completed task 3, do your best to *create your own concept map* to fit your problem.

A Concept Map is the map showing the inter-relationships between different concepts or ideas. An example below is the concept map of living things and some of their properties and examples.

```
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Consume</td>
<td>Energy</td>
<td>Reproduce</td>
<td>Perish</td>
</tr>
<tr>
<td>Does</td>
<td></td>
<td>Does</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Does</td>
<td></td>
</tr>
<tr>
<td>Animal</td>
<td></td>
<td>Example</td>
<td>Example</td>
</tr>
<tr>
<td>Have</td>
<td></td>
<td>of</td>
<td>of</td>
</tr>
<tr>
<td>Legs</td>
<td></td>
<td>Help</td>
<td>it</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Can</td>
<td>Move</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Can not</td>
<td></td>
</tr>
<tr>
<td>Plant</td>
<td></td>
<td>Have</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Leaves</td>
<td></td>
</tr>
</tbody>
</table>
```

**4.4.2 End-semester individual interview**

Interviews conducted at the end-point of the semester also included four problem tasks. For the final individual interview, students were not asked to complete a task regarding concept maps, but they were asked to complete a second text editing task. Tasks assigned for the second interview incorporated problem concepts most recently covered in the algebra-based physics course.
**Task 1 – Problem Posing**

Problem Posing was assigned as task 1 of the interview. Students were asked to create two separate problems for each scenario involving two constraining conditions.

**Figure 4.19 Problem posing task assigned for individual interview 2.**

You are provided two concept scenarios below. Create (make up) your own problems to fit each of the concept scenarios. Once you’ve made your problem, draw a diagram or picture to accompany it. You will NOT be asked to solve the problems you create.

**Concept Scenario 1:** Bernoulli’s principle and equation of continuity to determine the velocity of a fluid.

**Concept Scenario 2:** Angular momentum is conserved, angular velocity of an object increases.

**Task 2 – Physics Jeopardy**

Task 2 was a physics Jeopardy task. Students are given a solution or partial solution to a problem, and asked to identify a scenario that corresponds to solution.

**Figure 4.20 Physics Jeopardy task assigned for individual interview 2.**

The given information below is a worked-out solution to a Physics Problem. First identify what concepts are used to solve the problem. Then describe a real-life situation that best fits this solution.

\[ F_T = m_{hull}g - \rho_{water}V_{submerged}g \]

\[ F_T = m_{hull}g - \rho_{water} \frac{m_{hull}}{\rho_{hull}} g \]

\[ F_T = (1.8 \times 10^4 \text{kg})(9.8 \text{m/s}^2) - (1.00 \times 10^3 \text{kg/m}^3)(\frac{1.8 \times 10^4 \text{kg}}{7.8 \times 10^3 \text{kg/m}^3})(9.8 \text{m/s}^2) \]

Definition of the term “hull” – the frame or body of a large vehicle. Examples include ships, airships, submarines, and tanks.
**Task 3 – Text Editing**

Students were given a problem statement in task 3 and students were asked to find irrelevant or missing information. If none existed, students were asked to declare the problem sufficient.

Figure 4.21 Text editing task assigned for individual interview 2.

Please specify whether the problem statement below provides sufficient, missing, or irrelevant information for applying toward a solution.

A bullet is fired, moving horizontally with a velocity of 800 m/s before impacting a 2.50 kg block of wood which is suspended like a pendulum from a 3 m long rod. Assume the rod has negligible mass. As a result of the inelastic collision, the pendulum with the bullet stuck inside it will swing up to a maximum height. Determine the maximum height of the pendulum with the bullet stuck inside it.

**Task 4 – Case reuse text editing**

Task 4 would use the same context, but the information given and information missing was altered from task 3. The purpose of the second text editing task was to determine whether students could reuse information retained from task 3 to assist with task 4 more readily.

Figure 4.22 Text editing task assigned for individual interview 2.

Please specify whether the problem statement below provides sufficient, missing, or irrelevant information for applying toward a solution.

A 0.010 kg bullet is fired horizontally into a 2.50 kg block of wood which is suspended like a pendulum from a 3 m long rod. Assume the rod has negligible mass. As a result of the inelastic collision, the pendulum with the bullet stuck inside it will swing up to a maximum height. Determine the maximum height of the pendulum with the bullet stuck inside it.
4.5 Results - Individual Interview

A total of 19 students completed both individual interviews. Eleven students participated in the group learning interviews treatment and eight students sharing similar initial grade distributions were selected from the original volunteer list as a baseline group. Grade distributions were compared up to the first examination. Student performance on the individual interview task is described below. Participants in our group learning interviews and our baseline were assessed in the same way with exception to task 4 of individual interview 1. Task 4 required the creation of a concept map, and therefore would put our treatment group at a slight advantage. Students participating in the focus group learning interviews were assessed with regards to whether they created a concept map and by how much their map incorporated features of a previously seen structure map. Students in the baseline group were assessed as to whether a map was created. Data collected during the individual interviews were video and audio taped.

4.5.1 Individual Interview 1

Task 1 – Problem Posing

Students studied kinematics, forces, centripetal motion, work and energy, and momentum prior to participating in individual interview 1. Students were asked to pose a problem pertaining to a given visual queue for task 1 of individual interview 1. Figure 4.15 above shows the problem posing task given.

Concepts covered by posed problems: Student participants in both the treatment group and the baseline group focused their problem statements such that they were solvable using kinematics. Seven of 11 students participating in the FOGLI’s created kinematics problems. Six of 8 students in the baseline created kinematics problems. Other problem types were created, but they were often second and third problems created after being prompted by the interviewer to create more than one. Suitable momentum, angular momentum, and work-energy problems were also produced.

Level of real-world detail attributed to the problem statement: Twelve of 19 students participating in these interviews took extra care to either create problem statements that used correct notation and quantity size, or noted that it was important to
do so. Six of the 19 students were focus group learning interview participants. It was
difficult for these students to leave quantities as variables, but they would do so if they
were unable to come up with an appropriate quantity and the interviewer asked to move
on. All 19 students also preferred to make assumptions that were not real-world viable,
such as assuming the goalie would stand out of the way from the goal, the blockers on the
field would not move or raise their arms, or the blockers were all identical in height.
When students were asked to explain why their assumptions were necessary, student
responses were similar. Most indicated that the assumptions were necessary because the
problem would be too complex if information was given about each blockers’ reaction.

Student “You need to know how each blocker would react when the ball comes
by them, and it’s just easier to assume they don’t move…you could put
information in there if you wanted to, but then the kicker would have to
put spin on the ball in order to avoid all the blockers and I don’t know how
to calculate it.”

Two students indicated that the assumptions were close enough to an actual situation
where the blockers do not react to the ball movement quickly enough.

**Differences between the baseline and FOGLI participants:** The only noticeable
difference between the treatment groups and the baseline groups is between the number
of problems created for task 1. Students from the baseline group came up with two or
more problems on average, and students from the treatment came up with one on average.
Though more problems were created by the baseline, the problem sets tended to resemble
one another in context and concepts.

*Task 2 – Text Editing*

Task 2 required students to look at a problem statement and determine whether
the statement had missing or irrelevant information. Figure 4.16 displays the task given
for interview 1.

**Identification of Irrelevant information:** Two students, one participating in the
treatment and one in the baseline, were unable to identify the mass as being unnecessary,
and declared the problem as sufficient. Students that recognized mass as unnecessary
cited the range equation as an explanation as to why mass would simply cancel out. Most
students, when asked to further explain why mass cancels out, restated the range equation as a simple fact.

Interviewer “Why does the mass cancel out for this problem?”

Student “Because the equation contains no mass. This problem is solved using this equation.”

Interviewer “If you were to think about this situation conceptually, what does it mean for the mass to cancel out?”

Student “It means it’s not necessary for problems that use this equation.”

Three students, one from our treatment and two from the baseline, cited that gravity acts on all objects with the same constant acceleration and so mass is unimportant to the speed at which an object falls.

**Identification of missing information:** Two students, both from the control group, stated that the acceleration due to gravity was not given in the problem statement. The three students that explained how mass was independent of the speed at which objects fall also stipulated that they had to assume air resistance was not-existent or negligible. One of the three went on to say that the problem was either missing information regarding the air resistance or contained unnecessary information with regards to the mass, and therefore, there was more than one answer to this task.

**Differences between the baseline and FOGLI participants:** There were no discernable differences between the two groups. Both groups contained a majority of students capable of picking out the correct irrelevant information. The three students that were capable of thoughtfully explaining why the information was irrelevant were distributed between both groups.

**Task 3 – Physics Jeopardy**

Task 3 required students to look at a problem solution and create a scenario that best describes the solution. Figure 4.17 displays the physics Jeopardy task used during interview 1.

**Recognition of concepts:** All students recognized conservation of energy from the first two lines of the solution which describe the relationship between potential and kinetic
energy of an event at its initial and final condition using equations without numbers plugged in for variables.

**Recognition of quantities:** Fifteen students, eight from the treatment, correctly recognized the individual components in the first two equations of the solution as either gravitational potential or kinetic energy. Four students recognized the final line as final step for solving for the maximum height (the previous equation had numbers filled in and the height left as a variable). Eleven students, four from the baseline group, were unable to consistently recognize components of the solution while creating a situation. Some students recognized that the initial kinetic energy term in one moment, only to refer to it as a potential in the next.

**Recognition of scenario:** Three students were able to produce scenarios that were acceptable for the solution given. Two of these three students were a part of the baseline group.

**Following directions:** Though the task did not require students to create a full problem statement, only describe the scenario, all but two students wrote out a full statement with question. The misinterpretation of the directions given is not surprising. Intuitively a problem solution should have a corresponding problem.

**Differences between the baseline and FOGLI participants:** There were no noticeable differences between the baseline and FOGLI treatment students. Students’ difficulties with recognizing the individual components of given equations were not isolated to one group or the other.

**Task 4 – Creation of a Concept Map**

Task 4 requires students to create their own concept map that describes the previous scenario created in task 3. See Figure 4.18 for task 4 of individual interview 1.

**FOGLI participants map resemblance:** All students in the treatment group created structure maps, and nine students created maps that mimicked structure maps created in treatment. Two treatment group students created maps that mimicked the example given in the task instead. It was apparent from this task that students in the treatment groups were capable of at least remembering the general outline of the work and energy structure map. The two students from the treatments that created a concept map mimicking the
example, both apart of group 1, noted that their maps were not supposed to look like the structure map given in the focus group learning interviews.

G1S5 “I intentionally did this.”
Interviewer “You intentionally did ..what?”
G1S5 “Haha, yeah ok, I intentionally made my map look like the example…with words on the lines and the circles not being just the variables in the problems, but also being actions.”
Interviewer “Ok”
G1S5 “Is that alright? do you want me to make another one that looks more normal?”
Interviewer “Not unless you want to. There is no right or wrong here. Create the map that you think best represents this scenario.”

**Baseline participants map resemblance:** All but one student in the baseline group created maps that mimicked the example map given in the task. One student refused to create a concept map in the baseline group because they were asked to create concept maps in high school and simply refused to use or create them ever again.

**Differences between the baseline and FOGLI participants:** On average, baseline students create maps with four nodes and six associations between nodes. Students in the treatment group created maps with, on average, eight nodes and eight associations between nodes. Maps were neither assessed on correctness nor were they compared with expert maps. Since students in the baseline group were asked to create a concept map like the one given in the example, and the example did not look like the structure maps used in the treatments, it was not unreasonable that the final concept maps created by the baseline group would not resemble maps created by the treatment group students.

### 4.5.2 Individual Interview 2

**Task 1 – Problem Posing**

Students were asked to provide two separate scenarios pertaining to two separate concepts scenarios, each with multiple constraining conditions. See Figure 4.19 for task 1 of interview 2.
**Fulfillment of the conditions in concept scenario 1:** Sixteen of 19 students created a problem fulfilling only part of the scenario first, and then modifying it once the moderator pointed out the missing component. All sixteen students were able to successfully modify their original problem to include second conditions. Three students total, all a part of the treatment groups, were unable to create a scenario using both conditions.

**Fulfillment of the conditions in concept scenario 2:** Many students were capable of providing an appropriate scenario using only one of two scenario constraints in task 1, part 2. Eight students, four from the treatment group and three from group 1 of the treatment groups, could create problems that would have increasing angular velocity, but did not understand what it meant to have conservation of angular momentum. These students created problems where external forces would be required for the increase in angular velocity, but referred to the general scenario as being ‘angularly conserved.’ Students that had trouble with conservation of momentum would often cite the net torque as equal to zero as an explanation for conserved angular momentum, without explaining what that meant for their particular scenario. Eight other students, five from the treatment groups, would create problems where angular momentum was conserved, but could not figure out how one could increase angular velocity with conserved angular momentum. Of the three students capable of completing the full task correctly, two of these three were participants in our focus group learning interviews. All three of these students needed the textbook to cite a specific example that they had talked about in class.

**Differences between the baseline and FOGLI participants:** There were no differences between the two groups with regards to correctness or completeness. Common student difficulties with this task were equally distributed among all individual interview participants.

**Task 2 – Physics Jeopardy**

For task 2, students were given a partial solution to a problem, and were asked to create a scenario that best describes the solution. See figure 4.20 for the physics jeopardy task used during individual interview 2.
Recognition of a specific and necessary surface feature: All students recognized that a ship or submarine must be involved in the problem from the subscripts and hint given. It is apparent that unlike the previous physics Jeopardy task in individual interview 1, the more specific context seems to cue a different problem solving scheme. Students spent less focus on taking apart mathematical representations and spent more time attempting to infer a situation from their prior knowledge about boats and submarines. This became more apparent as students’ described the possible direction of the tension force given in the problem solution.

Recognition of quantities: All but 1 student, in treatment group 1, recognized $F_T$ as a tension force. Student G1S2 recognized $F_T$ as a torque. Seventeen of 19 students recognized the individual components of the tension force as the buoyant force and the weight force.

Recognition of direction: Seventeen of 19 students were unable to determine the direction of the tension force, and guessed a direction based upon reasonable assumptions made about boats or submarines. Twelve of these students, five of which were apart of the treatment groups, assumed that the boats have anchors, while five students including three treatment participants, assumed that boats must be placed in water using a cable.

Student “It’s the force due a dropped anchor.”

Five of the guesses were correct and four of these five guesses resulted from flipping around the text and looking at examples, while the fifth resulted from a previous problem example the student remembered having seen. Two students, one from the treatment groups and the other from the baseline group, were capable of creating a scenario that fit the solution without discussion of tension direction or use of the textbook as a resource.

Differences between the baseline and FOGLI participants: Again, students from the treatments did not perform differently from the baseline students on this task.

Task 3 – Text Editing

Task 3 was a text editing task, that included one irrelevant piece of information, and was missing another. Figure 4.21 displays the text editing task used for individual interview 2.
Identification of problem statement as sufficient: All student participants but one in our treatment group marked the problem statement as sufficient. The one student that did not declare the problem as sufficient, correctly determined that the initial velocity was missing and that the length of the rod was irrelevant.

Task 4 - Case Reuse Text Editing

Task 4 was also a text editing task. The irrelevant information remained the same as task 3, but the missing information from task 3 was now provided and a different quantity previously given in task 3 was removed. This task may be viewed in Figure 4.22 above.

Identification of missing information: Only two students were unable to identify the velocity as missing. All other students were able to recognize the missing velocity by what they had seen from the previous task. Of the two students unable to identify the missing velocity, one from the treatment groups and the other from the baseline group, neither could identify missing information even when told they could look at both task 3 and 4 side by side.

Identification of irrelevant information: Sixteen of the 19 students were unable to recognize the rod length as being unnecessary. Two of three students that recognized the rod length as being irrelevant were in the baseline group. The third student was from the treatment group, and recognized the same irrelevant information given in the previous task. All students that recognized the rod length as being unnecessary would ultimately be able to provide explanation as to why the rod length conceptually was unnecessary.

Student “If the rod was shorter or whatever, I would still reach the same height because I still had the same amount of stored energy.”

Two students, one from the treatment groups and one from the baseline, had sufficient time to modify the problem such that the rod length became important, a task which was not officially asked, but informally requested given enough time.

4.5.3 Summary of Individual Interview results

All students participating in the individual interviews completed the tasks required. We were able to address each of our original research questions from the
written data and video recordings collected in each interview. Each of our original research questions will be answered here.

**How complete were students’ responses to problem tasks assigned in the individual interviews?**

All of our 19 students completed the three non-traditional tasks for individual interview 1. Eighteen of 19 students completed the fourth and final task which required students to create a concept map. One student from the baseline group refused to complete the task because she felt concept maps were not a constructive means for representing concepts.

All of our 19 students completed the four non-traditional tasks for interview 2.

**How correct were students’ responses to problem tasks assigned in individual interview 1?**

The problem posing task assigned was measured as correct if the participant was capable of articulating at least one problem statement that incorporated the scenario given in the presented picture. All students were capable of creating at least one distinct problem, and many were capable of recognizing the importance of approximating the real world scenario by using proper notation and correct quantity magnitudes.

For the text editing task, 17 of 19 students identified the correct irrelevant information. Three of the 17 were able to correctly explain why the irrelevant information based on conceptual reasoning. Others simply stated an equation as evidence that the quantity would eventually ‘drop away.’

All students were able to recognize that the physics Jeopardy solution involved conservation of energy. Four students were able to recognize that the problem asked them to solve for a height. Three of those four students were capable of producing a scenario that matched the solution given.

For the concept map task, all but one student successfully created a concept map. They were not analyzed based on correctness. Rather they were analyzed to determine whether the quantities and concepts present in each representation aligned with the appropriate scenario.

**How correct were students’ responses to problem tasks assigned in individual interview 2?**
For the problem posing task, 16 of 19 students successfully created problems for scenario. All sixteen students required prompting to include both conditions in their final problem. Three of the 19 students were capable of successfully creating problems for scenario 2. The textbook was used to reference an example used previously by the three students. Sixteen students (9 participating in our FOGLI’s) were capable of creating a problem that would incorporate only one of the two conditions for scenario two.

For the physics Jeopardy task, seven students, 4 participating in our FOGLI’s, were able to create an appropriate problem scenario matching the solution presented. Ten other students were able to recognize that the scenario would need to include some vertical tension on a watercraft.

For the first text editing task, only one focus group learning interview participant recognized both the irrelevant and missing information given/not given in the problem statement. The 18 remaining students determined that the problem statement was sufficient.

Seventeen of 19 students were able to recognize the initial velocity as missing required information for the second text editing problem. Only three of 19 students were able to recognize the rod length as irrelevant information.

**How did our cohort compare with the baseline group in the individual interviews?**

There were no differences between the baseline group and focus group learning interview participants for any of the non-traditional problem tasks comprising individual interview 1. There were no differences between the baseline group and focus group learning interview participants for any of the non-traditional problem tasks in individual interview 2. Differences between the group were seen only in the final task of interview 1, where students were asked to create a concept map for a work and energy problem. Students participating in the focus group learning interviews were more adept at creating concept maps that looked like a previously used work and energy structure map. Students in the baseline group were capable of creating many variations of concept maps loosely based off the example given in the task statement. The differences between the concept maps created by the baseline group and the focus group learning interview participants were a measure of problem solving performance. The significance is only
that our focus group learning interview participants were capable of recalling the structure maps from memory.

4.6 Summary of Results

We answered our research questions for this phase of the research project.

How do students use structure maps while problem solving?

Our results indicate that students have difficulty using the quantities in a structure map to solve problems if they are not provided explicit equations. Students appear to like the work and energy map, but we do not know whether it was the map’s temporal symmetry or whether the topic is just better understood by these students. A map with equations in the nodes, like the vibrations and waves map, enabled students’ to recognize connections between individual quantities inside equations and was found useful by the students.

Feedback from students led us to change the structure maps from those in which the nodes contained physical quantities to those in which the nodes contained equations, with arrows showing how quantities between equations were related. These kinds of maps appeared to provide students a pathway to connect the equations and were found to be useful by the students in problem solving.

How useful do students perceive the structure maps?

While students perceived several maps to be useful in problem solving, there is no evidence that these kinds of maps facilitate expert-like problem solving strategies from problem performance during the focus group learning interviews or during the individual interviews. Student participants become more efficient at completing problem solutions and using structure maps during interviews, but it is difficult to determine whether that is due the temporal and spatial symmetry differences between maps, the concepts expressed in the problems and the structure maps, or the change in problem complexity resulting from a change in concepts.

How does student performance change, if at all, over the semester using non-traditional problem solving tasks as an assessment aimed at measuring students’ efficiency of information processing?
Students continue to use novice-like, plug and chug strategies when provided with the revised map, and there are no definitive differences associated with problem performance between our focus group learning interview participants and the baseline group. Therefore, this study provides no evidence that structure maps, as used here facilitate expert-like problem solving in physics.

**How does our cohort compare with the baseline group of students on non-traditional problem solving tasks?**

There is no noticeable difference between the treatment group participants and the baseline group in the individual interviews. Completeness and correctness of solutions to given tasks were variant and possibly dependent on the type of task, the concepts covered, and the context chosen for distinct problems. It would be nearly impossible to compare each non-traditional task between interviews for this reason. All groups remained similar in grade distribution for the class as well.

### 4.7 Limitations and Future Work

The results of this study are clearly not promising. One possible reason is that the skills for using structure maps in expert-like problem solving need to be developed over long periods of time. Our study was clearly limited in scope and the types of structure maps used. Future work would require a larger experimental sample and a larger variation of structure maps. It might also be interesting to note whether order of principles has an effect on student preference such as introduction of the force and centripetal acceleration problems again at the end of the semester treatment. This suggestion would be feasible if students were given a comprehensive exam as the review of previous material would be perceived as helpful. However, the focus for this project is to elicit higher development of students’ organization of retained knowledge by facilitating the case reuse strategy. The structure maps could be redeveloped to better model organization of student knowledge while still incorporating some elements of student suggestions, but that would require its own study. It was determined that for the remainder of this project, time would be better spent if the project focused on explicitly observing and facilitating case reuse by asking students to contrast and compare problem
pairs. The pairs could emphasize similarities in principles and concepts while varying the physical surface features sometimes misconstrued as being concept and principle dependent (i.e., a block on an incline, a mass attached to a pulley). This facilitator is also more readily implemented in an actual classroom.
CHAPTER 5 - PHASE II – Focus Group Learning Interviews to Facilitate Case Reuse in Problem Solving

5.1 Introduction

This chapter describes the second implementation of group learning interviews with an algebra-based student cohort. The group learning interviews focused on assisting students’ organization and processing of knowledge learned while students encoded information on how similar problems are solved. Jonassen (Jonassen, 2006) suggests that students assemble and/or modify schemata (mental representations) by analyzing a worked example. These representations include information regarding the problem type, structural elements (velocity, distance, etc.), situations in which the problem occurs (monkey in a tree, inclined plane, etc.), and the processing operations necessary for acquiring a solution. Schemata are retrieved as learners work solutions to new similar problems (case reuse).

Previous phases of our study used pre-constructed structure maps as a facilitating strategy through which students could develop problem schemata. A structure map, like a concept map, is a visual organization of physical quantities. They were designed by physics experts to emulate students’ problem schemata as theoretically constructed by Dedre Gentner (Gentner, 1983). The structure maps were chosen because they offered a visual representation of the associations between quantities and students could readily compare similarities and differences between problems using the maps. Though the structure maps were well-liked by students after altering the original expert map design to better fit students’ proximal comfort zone, the process of learning how to use the structure maps was time consuming, and the effect on case reuse was not explicitly measured quantifiably or qualitatively.

Our goal was to facilitate the development of conceptual problem schemata during problem solving using case-reuse strategies. To achieve this goal for phase II, we conducted group learning interviews with a cohort group of students enrolled in an algebra-based physics course. Participants included 10 students that were randomly selected from 46 volunteers. The participants met in a single group a total of eight times.
during the semester for what we called focus group learning interviews. Each of the eight focus group learning interview sessions was 75 minutes long. Additionally, the participants met with the moderator individually twice during the semester. The individual interviews were conducted at the mid and end points of the semester. Each individual interview was 50 minutes long. Figure 5.1 shows the research methodology on a timeline beginning from left at the start of the semester to the right at the end of the semester.

**Figure 5.1 Research design timeline.**

The focus of these interviews was restructured from previous phases to align with *explicit* contrasting of problems that help students focus on deep structural properties of a problem rather than surface differences. Previous research suggests that learners fail to recall examples or schema appropriately because their retrieval is based upon similarity of objects (car, boat, incline, etc.) between examples, not their structural (conservation of energy, momentum) characteristics (Catrambone and Holyoak, 1989; Reed and Bolstad, 1991). Catrambone and Holyoak also suggest that generalization improves when problems *emphasize* structural features shared with a similar example. Tasks assigned to students during these new focus group learning interviews would include explicit contrast and comparison of problems emphasizing structural features shared with a similar example. We assessed the impact of our intervention on students’ problem schema using problems posed by students during our Focus Group Learning Interviews, non-traditional problems on exams, and non-traditional problems given during the individual interviews. At the end of this chapter, we look to answer three research questions:
• How do students perceive the usefulness of an example problem given a new problem based on the same physical principle but lacking similar objects and/or orientations?
• Given problems that are deep-structure similar and surface feature different, what similarities and differences do students offer as important to the problem solutions?
• How does students’ emphasis on similarities and differences change given problem pairs with varying deep structure and surface feature similarities / differences?
• How does students’ performance on problem solving differ with respect to the rest of the class?

5.2 Screening interviews

Screening interviews, each lasting about 20-30 minutes were conducted with 21 participants that were selected from a pool of 46 volunteers. Each student was paid $8 for participating in the screening interview.

The main purpose of the screening interviews was to gain insights about how students solved problems and whether or not they worked with others. Students were asked about the prior physics classes they had taken, including in college and high school. See Appendix C-1 for screening interview protocol. They were asked about their interest in the current physics class that they were taking as well as why they were taking it. Thus, we wanted to screen for students who were not apathetic toward the class or were very likely to drop out in the middle of the study.

We were most interested in selecting students who would be amenable to learning how to solve problems by looking at solved examples and also those who were comfortable working with others, since the group learning interviews were an interactive environment and we wanted to ensure that students who were selected in our study would be comfortable participating in it. To screen for these attributes, we asked students about their study habits, especially how they went about solving problems. A significant aspect of the study was case-reuse - whether they used solved examples, and if so how. Based on our previous phase (Mateycik, Jonassen et al., 2008) we had seen that students tended to overly rely on using equations. We asked students if and how they would use
equations in solving problems. Finally, we also asked students whether or not they found the textbook to be useful and if so, in what ways was the textbook useful to them.

### 5.2.1 Results from screening interviews

Most of the 21 interviewees who participated in the screening interviews had taken physics in high school. They were primarily construction science and animal science majors and were currently enrolled in the only sequence of physics classes required for their respective majors. See the table below for complete listings of participants demographics by their code ID.

**Table 5.1 Demographics for Phase II participants.**

<table>
<thead>
<tr>
<th>Code ID</th>
<th>Selected for Treatment</th>
<th>Major</th>
<th>Previous Physics Classes</th>
<th>M/F</th>
<th>Class year</th>
</tr>
</thead>
<tbody>
<tr>
<td>AM</td>
<td>Y</td>
<td>Animal Science/Pre-Vetinarian</td>
<td>None</td>
<td>F</td>
<td>1 yr</td>
</tr>
<tr>
<td>AR</td>
<td>Y</td>
<td>Kinesiology</td>
<td>High School Physics</td>
<td>F</td>
<td>3 yrs</td>
</tr>
<tr>
<td>CF</td>
<td>Y</td>
<td>Construction Science</td>
<td>High School Physics</td>
<td>M</td>
<td>2 yrs</td>
</tr>
<tr>
<td>DM</td>
<td>Y</td>
<td>Animal Science/Pre-Vetinarian</td>
<td>High School Physics</td>
<td>F</td>
<td>2 yrs</td>
</tr>
<tr>
<td>EJ</td>
<td>Y</td>
<td>Biotechnology/Pre-Vetinarian</td>
<td>High School Physics</td>
<td>F</td>
<td>3 yrs</td>
</tr>
<tr>
<td>MD</td>
<td>Y</td>
<td>Biology</td>
<td>Physical World</td>
<td>F</td>
<td>4 yrs</td>
</tr>
<tr>
<td>MM</td>
<td>Y</td>
<td>Construction Science/Management</td>
<td>None</td>
<td>M</td>
<td>2 yrs</td>
</tr>
<tr>
<td>MR</td>
<td>Y</td>
<td>Construction Science/Physics</td>
<td>High School Physics</td>
<td>M</td>
<td>3 yrs</td>
</tr>
<tr>
<td>MS</td>
<td>Y</td>
<td>Food Science/Nutritional Sciences</td>
<td>High School Physics</td>
<td>F</td>
<td>3 yrs</td>
</tr>
<tr>
<td>SS</td>
<td>Y</td>
<td>Lifescience</td>
<td>Descriptive Physics</td>
<td>F</td>
<td>4 yrs</td>
</tr>
<tr>
<td>SB</td>
<td>N</td>
<td>Animal Science/Pre-Vetinarian</td>
<td>None/None</td>
<td>F</td>
<td>2 yrs</td>
</tr>
<tr>
<td>MC</td>
<td>N</td>
<td>Biology</td>
<td>None/None</td>
<td>F</td>
<td>2 yrs</td>
</tr>
<tr>
<td>LC</td>
<td>N</td>
<td>Biology</td>
<td>Physics 1 Honors</td>
<td>F</td>
<td>1 yr</td>
</tr>
<tr>
<td>TD</td>
<td>N</td>
<td>Construction Science</td>
<td>High School Physics</td>
<td>M</td>
<td>3 yrs</td>
</tr>
<tr>
<td>BG</td>
<td>N</td>
<td>Nutrition/Pre-Med</td>
<td>None/None</td>
<td>F</td>
<td>3 yrs</td>
</tr>
<tr>
<td>YH</td>
<td>N</td>
<td>Biology</td>
<td>International - 2 yrs of Physics</td>
<td>M</td>
<td>2 yrs</td>
</tr>
<tr>
<td>WJ</td>
<td>N</td>
<td>Biology/Gerontology/Spanish</td>
<td>AP Physics</td>
<td>F</td>
<td>3 yrs</td>
</tr>
<tr>
<td>JO</td>
<td>N</td>
<td>Biology</td>
<td>High School Physics</td>
<td>M</td>
<td>1 yr</td>
</tr>
<tr>
<td>SL</td>
<td>N</td>
<td>Animal Science/Pre-Vetinarian</td>
<td>High School Physics</td>
<td>F</td>
<td>3 yrs</td>
</tr>
</tbody>
</table>
This interview was conducted in the first two weeks of the semester. The views expressed by students were therefore primarily based on their experiences in high school physics classes they had taken and not necessarily based on their experiences in their current college physics class.

When asked about their study habits and use of the textbook, 12 of the students felt that the book was well written and easy to read. The other nine students felt that the book was difficult to read and limited the extent to which they would read the text.

Student “The physics book is really kinda easy because it tells you like what section you'd draw from like let's say section 2.5 to the following question, section 2.6, and you know some books are like chapter 2 you gotta look over the whole chapter, but the physics book that we got is kinda easy so, i'd just look at the homework problem, i'ld just look back, just to find if it was related, example problem, just go by that”

The problem solving procedures described by students seemed to be similar to each other. All 21 screening interview students stated they would first read the problem, and pick out information that was given and asked for in the problem. Nineteen of 21 students also stated they would solve the problem using one or a set of formulas that contained quantities that were both given and asked for in the problem.

The other three of the original 21 stated that they often needed to think carefully about what to do because they could not just plug in numbers.

Student “I can understand [the instructor] when he's talking about it and understand it when the book says it, but then if I know if I have to do a problem about [a given topic] I feel like "ok where do I start?" because I've always buttoned up way more on the word problems. I mean I understand it when the teacher saying, when someone is walking me threw,
and this is how you do it, and I'm like ok! I understand… but I don't know
when I start, I can’t plug in to equations he used.”

Sixteen of the 21 students also stated they would use solved examples to help them with
homework problems.

Student “I can find for the most part an example… is pretty much parallel to the
homework problem so, I like the example a lot for the formulas.”

Twenty of 21 students also reported that they would read or reread the relevant section of
the textbook as well as find solved examples from the section that they believed would help them solve problems. When these students were asked specifically about their
current physics homework, all of the students found the problems on the homework
assignments for the first and second week to be quite easy to complete.

Student “I read the textbook before every class so you kinda know what's going
on, but right now we're doing stuff that I knew from my past physics class,
so I'll know more if it is helpful when we get to new stuff that I don't
remember very well. It kinda helps explain something if you don't
understand it in the lecture, cause you can't ask many questions in the
lecture. I don't use [example in the text] for this class but i have for my
first [physics course]… I remember we've done something like [our
homework problems] before, so in a way I do I guess.”

Based on the screening interviews, we invited 10 students to participate in the
focus group learning interviews. All students that participated in the screening interviews
appeared to be interested in and looking forward to the class. Only one student voiced an
opinion that they might drop the class early. The ten students selected were chosen
randomly from the remaining 20 students not likely to drop. Students that mentioned that
they often worked with a study partner while solving homework problems were also
preferred in the selection, though only two of the 21 stated that they preferred working
alone.
5.3  Focus Group learning Interviews

A total of 10 volunteers were selected from our screening interviews. They were invited to participate in the focus group learning interview sessions. A total of eight focus group learning interview sessions were held during the semester – about one per week, except on weeks when students had exams and other commitments. The topics addressed in each group learning interview are also listed in Figure 5.1. These topics cover the typical topics that are covered in a first semester algebra-based physics course. During each focus group learning interview session a moderator would hand out a pair of problems for students to work on. These problems were labeled problem A and problem B. Participants were paired together such that one would be asked to work on problem A while the other worked on problem B. The problems shared deep structure similarities but had surface differences. After students had solved these problems they were asked to discuss their solutions with their partner briefly and discuss the similarities and differences between each of the problems. Students were also asked to work with their partner to create their own problem which uses some elements from both problem A and problem B. If time was allowed, students were also asked to switch problems with their neighboring group and to solve the problem if possible. Some groups were able to get through these group tasks faster than others, and students were asked to make sure they completed tasks related to the similarities and differences prior to moving on to the problem posing task.

5.3.1  Focus Group Interviews-Week 1

**Methodology**

An example of the problem pairs used in the first interview is shown in Figure 5.2. Both of the problems present contrasting cases and are focused on the same physical principle (Newton’s II Law), but have many surface differences such as vertical versus horizontal orientation and the different kinds of objects (blocks versus train cars) in the two problems.
Results from Week 1

In the first week of the semester, students often struggled with completing their individual problems. They had difficulties solving the problem and therefore did not have the time to engage in problem comparisons and discussions. Students are asked to compare their solutions to Problems A and B and come up with a new problem involving elements of the previous problems. General discussion as a group did not cover student created problems.

Principles – With exception to one student, everyone identified the concept of Newton’s 2nd law without difficulty. Student AL actually used Newton’s 2nd law while solving the problem, but identified it incorrectly. Students that did not identify Newton’s 2nd law applied F=ma in their solution. Students were asked to elaborate as to why they felt these problems involved Newton’s second law, but students became agitated and cited the book chapter as being obvious evidence of the ‘principle involved.’

Student MD “Because this is what we did in class…and its in our chapter section.”
Student EJ “Yeah, Fran”

**Text Editing** – General discussion included talk about how to extract important information from a problem. Students readily identified the mass of the cars as being unimportant, but discussion was not focused on why this was true. Confusion included but was not limited to the extraction of missing, but implied information. Eight students were unsure of how they were to know that they needed to extract information about the acceleration from the problem.

Student EJ “Where do you come up with acceleration is zero?”
Two students in the group answered that it has a constant velocity. Student AM states that “You don’t actually need the number for velocity, you just need to know that it is constant because you only use acceleration equal to zero.”

Students also voiced concern over the force of tension remaining over a constant velocity. Students unanimously decided that tension must change as the speed changes. Four students felt that the problem with the train must include friction between the cars and the track, while the rest of the cohort felt friction must be ignored. Students that wanted to include friction felt that you needed friction to keep tension between the cars, otherwise the cars moving at a constant velocity would act just like two cars that were stationary.

Student AM “Stationary cars must not have tension between them. Since the speed is constant, acceleration is zero, and the situation is just like as if the cars are sitting there because constant speed and zero speed both have zero acceleration.”

Two students also voiced a concern over how to draw the diagrams.

Student MM “Do you draw your diagram at the coupling between cars or at the center of the car?”

One students asked if air friction might effect the solution to the problem. Two others immediately answered that air friction is not apart of the problem because they have not incorporated air friction into problems in class yet.

**Similarities and Differences** – Students identified the tension or force is being applied between two different objects as being the primary similarity between the two problems.
Students identified the direction of the forces as being the primary difference. Problem A included vertical forces while problem B included horizontal forces.

**Problem Posing** - Students did not have sufficient time to trade problems with one another. Four separate problems are posed by the student groups. Two problems incorporated boxes (instead of train cars) into moving train type problems. One problem built from the original train problem replaces cars with boxes, constant velocity with a constant acceleration, and makes the assumption that the reader is familiar with problem B. The second box problem was much more unique in that it hung the interconnected boxes and asked for the tension between blocks 1 and 2 from the static system. This problem is solvable. Another problem built from the original train problem asked for the full weight of a car and also replaced the constant velocity with a constant acceleration. This problem was under-specified for part (b). The final problem created involved boxes being pulled off a table, which incorporated facial features characteristic of both problem A and B. Again the problem had a constant acceleration and not a constant velocity. This problem was also underspecified. The problems posed by students during this interview are in Appendix C-2 under the weekly summaries. Figure 5.3 is an example of an underspecified problem described directly above.
Students appeared to be having difficulty distinguishing acceleration and velocity during the focus group learning interview, and it was difficult for students to communicate the meaning of an acceleration or a tension force while asked to explain their problems.

5.3.2 Focus Group Interviews-Week 2

Methodology

To alleviate difficulties that arose from our time constraints, in the second week we introduced specific stopping points in the process at which students were asked to stop, signal to the facilitators and check with them about their progress in solving the problem. The problems given for week two are presented in Figure 5.4. The problems covered the topic of forces, and though both problems involved inclines, problem B was significantly different in terms of surface features with the inclusion of a pulley system. Figure 5.5 shows an example of the stopping points introduced in the new protocol.
Figure 5.4 Example of contrasting cases presented to students in week 2.

**Problem A**

Two boxes, $m_1=1.0 \text{ kg}$ with a coefficient of kinetic friction $\mu_{k1} = 0.20$, and $m_2=2.0 \text{ kg}$ with a coefficient of kinetic friction $\mu_{k2} = 0.10$, are placed on a $m_3 = 15.0 \text{ kg}$ plane inclined at $\theta=30^\circ$. A taut string is connected to the boxes. At the instant shown, block $m_2$ is moving downward at a speed of 0.05 m/s and is 1 meter farther down the slope than block $m_1$. What is the acceleration of each block??

![Diagram of Problem A](image)

**Problem B**

A box $m_1 = 28.0 \text{ kg}$, lying on a drafting table inclined at $\theta=30^\circ$, is connected to another box $m_2 = 14.0 \text{ kg}$ by a cord running over a frictionless pulley. The coefficient of static friction between the table and the block $m_1$ is $\mu_{s1} = 0.450$, and the coefficient of kinetic friction between the table and the block $m_1$ is $\mu_{k1} = 0.320$. At the instant shown, block $m_2$ is moving downward at a speed of 0.05 m/s and is 1 meter above ground level. What is the acceleration of each block??

![Diagram of Problem B](image)
Results from Week 2

Students were having difficulty solving the force problems given in this session. One of the moderators had to stop everyone at 40:00 minutes into the focus group to explain how each person should begin solving the problems. Students watched the moderator create a force diagram for problem A. The moderator continued for another 20 minutes to fully explain the solution to the problem. This left very little time for open discussion among the students.

Principles – Data collected from the worksheets suggests that students are capable of identifying the concepts involved in the problem as Newton’s second law. Due to the limited time left for conversation, the moderators were unable to determine how students determined the concept involved, though it should be noted that students used their textbook heavily during this second focus group learning interview.

Text Editing – Data collected from the worksheets also suggest that students given the mass of the plane were capable of identifying such information as irrelevant, but not necessarily capable of identifying other irrelevant information also contained in the same problem. Most students given the static friction as irrelevant were capable of identifying such information, but also did not mention other irrelevant features. The irrelevant information that used the same units as information that was relevant was less likely to be identified as irrelevant. (e.g., the length of the ramp has the same units as the height of the ramp.)
**Similarities and Differences** – Students were unable to complete the task of solving their problems, but they were asked to weigh in on the similarities and differences between the problems with their partners.

Students identified the following differences: There is no friction acting on second block. There is no normal force acting on second block. There is a pulley in one problem, but not in the other.

- Student MM “My problem, B, was way more challenging I think because there were objects moving in different directions. like, there was one on an incline and the other was just falling vertical, and they are attached to each other with a pulley.”
- Student JL “Problem A was harder though because both of our blocks were on an incline, and the inclines are always harder.”
- Student AM “Well, inclines require more math. I don’t think they are harder, you just have to break up the components of each block into x and y. I don’t know, I think they are both difficult.”

Students that were assigned problem A felt their problem was more difficult. Students that were assigned problem B felt their problem was more difficult. Students assigned problem B felt the pulley made their problem more complicated. They also felt that the problem was more difficult because the coordinate system was different between the first block and the second block. The second block was being pulled at an angle.

**Problem Posing** - Students did not have enough time during the week 2 focus group learning interview to complete the problem posing task.

### 5.3.3 Focus Group Interviews-Week 3

**Methodology**

In the third week, we provided more procedural scaffolding in the form of a more detailed stepwise process of how to solve the problem as shown in Figure 5.7. The purpose of the procedural scaffolding was to decrease the cognitive load of the students in following the particular steps to solve the problem, so that they would be able to attend to the conceptual aspects of the problem such as reflecting on the underlying principles, similarities, and differences between problems. Figure 5.6 includes problems given to
students during the focus group learning interviews. The topic covered was rotational motion using forces. Problem A and B shared similar features such as tables, but the number of objects in motion and the types of objects were different.

Figure 5.6 Example of contrasting cases presented to students in week 3.

**Problem A**
A puck of mass \( m = 1.50 \text{ kg} \) slides in a circle on a frictionless table while attached to a hanging cylinder of mass \( M = 2.50 \text{ kg} \) by a cord through a hole in the table. The distance from the puck to the hole is 0.02 m and 0.04 m from the hole to the cylinder. What frequency is required to keep the cylinder at rest?

**Step 1:** Identify and interpret the principles involved in the problem.

**Problem B**
Two blocks of masses, \( m_1 = 2.0 \text{ kg} \) and \( m_2 = 2.5 \text{ kg} \), are connected to each other and to a central 3 m high post by cords. Blocks 1 and 2 rotate about the post at a frequency 10 rev/s on a frictionless horizontal surface with distances of 1.0 m and 1.5 m from the post, respectively. Assuming the cords are connected to the post and both masses such that they are under tension in directions only parallel to the horizontal surface, what is the tension for each segment of the cord?

**Step 1:** Identify and interpret the principles involved in the problem.
Our experiences in the first three weeks of the semester taught us that students had difficulties in solving the problems. They had no time to reflect on the problem principles, similarities, and differences.

Results from Week 3

During week 3, students are asked to work through a step-by-step procedure in order to eliminate procedural difficulties related to the diagrams and problem solving that students had last week. Students voiced concern over the material covered in class. They spent over half the time looking for help in the textbook. They also spent a lot of that time complaining that they don’t like multiple choice exams.

“You get out from the test and your like, ‘yes!’ and you get it back and it’s like ‘Uggh’”

After 22:00 minutes, people were still struggling to get through problem. Off topic conversations continued to thrive and students could be heard giving up on the problem.

Student EJ “I don’t know”
Student JL “I don’t know how to do that”

Student AM “I wish we didn’t have a test coming up on this.”

At 33:08, one of the moderators stopped the students and began to run through the solution to Problem A. At 50:43, the moderator completed their run through of the problem solution.

During this point, students found a mathematical error in the solution to the problem where a value should have been squared, but was not. This resolution of a mathematical conflict seemed to entice students to further discuss the problems. The rest of our time was spent asking students to talk with one another about their solutions. Time ran out before students could share their solution or work on their own problems.

**Principles** – Students identified the ‘principles involved’ as forces and circular motion. Students briefly described the problem scenarios as congruent with the material covered most recently in class, but also stated they were unhappy with their understanding of the material.

Student AM “[Instructor] tried to go through this with us in class, and I thought I understood it, but we haven’t really done any problems like this before...[student goes off onto a tangent related to her disliking of the instructor]”

[Several students verbally agree with students AM]

Student MS “can you show us how to do this? I think we’re all lost.”

**Text Editing** – The students given problem B identified the height of the pole as irrelevant information rather quickly. All but two students given problem A were able to identify a length between the hole and the cylinder as being irrelevant. Two students present did not complete the worksheet. It was difficult to determine the reasoning for this selection of irrelevant information as this discussion was not included in this interview.

**Similarities and Differences** – Similarities and differences between the problems also remained unknown as students left this part of the worksheet blank, and the discussion never got that far.

**Problem Posing** - Students did not have enough time during the week 3 focus group learning interview to complete the problem posing task.
5.3.4 Focus Group Interviews-Week 4

Methodology

In week 4 students were given a worked example problem at the beginning of the group meeting. Research has shown that providing students with appropriate worked examples can facilitate problem solving (Ward and Sweller, 1990). The worked example problem, or problem C, included a full solution and was available as a resource for students to use while solving their own respective problems. Problem C would be deep structure similar to both problem A and problem B. Figure 5.8 shows an example of the solved example (Problem C) followed the contrasting pairs (Problem A and Problem B) in Figure 5.9. The facilitators went over Problem C in the first 15 minutes and then asked students to solve Problems A and B. All problems given for this week were work-energy problems.
Figure 5.8 The solved example problem C that was presented in week 4.

**Problem C**
Joshua pushes a 12.5 kg box along a flat horizontal table applying an average force of 39.0 N. The box starts at rest and reaches a velocity of 12.0 m/s. Neglecting friction, how far did Joshua push the box?

**SOLUTION**

We may express the work done by Joshua on the box in terms of the force applied and the distance covered while Joshua applied the force. We know that since the box will be moving in the same direction as the force applied, the angle between the direction of force and direction of motion is 0 degrees. Thus, the work done is given by:

\[ W = F_d \cos(0) = F_d \]

\[ \Rightarrow W = (39.0 N)d \]

We also know that the work done on the box must be equal to the change in kinetic energy. Since the box was at rest initially our initial kinetic energy will be zero.

\[ W = K.E_2 - K.E_1 = \frac{mv_2^2}{2} - \frac{mv_1^2}{2} \]

\[ \Rightarrow W = \frac{mv_2^2}{2} \]

Finally, we have expressed the work done on the box using two different equations above. We may set both expressions for Work equal to one another:

\[ W = (39.0 N)d = 133.2 J \]

\[ \Rightarrow d = \frac{133.2 J}{39.0 N} = 3.42 m \]

\[ \Rightarrow W = \frac{(12.5 kg)(12.0 m/s)^2}{2} = 133.2 J \]

Figure 5.9 Contrasting cases presented to students in week 4.

**Problem A**
A 0.10 kg arrow is fired from a bow. The bow is pulled back a distance of 0.8 m so that the arrow is released with a speed of 50 m/s as it leaves the bow. The arrow travels 25.0 m before hitting its target. What is the average force exerted on the arrow by the bowstring?

**Problem B**
A Yankee batter hits a 0.14 kg baseball sending it off into left field, 40 m away from the batter's box. The baseball lands in a Royals fielder's glove, exerting an average force of 300 N, moving the glove backward 0.25 m before coming to rest. What is the speed of the ball just before it is caught?

Based on our own observations of student performance in week 4 we realized that the protocol we had developed was successful in enabling students to work through the
problems without significant barriers. The solved example (Problem C) gave students adequate scaffolding to complete the unsolved problems (Problem A and Problem B). We asked students during week 4 to specifically score the usefulness of Problem C in solving Problem A and B. Figure 5.10 displays the worksheet questions pertaining to the usability and ranking of similarities and differences.

**Figure 5.10 Usability rating task for week 4.**

![Step 7: Discuss any similarities and differences that your problem has with your partners’ problem. Discuss the usefulness of Problem C while working on your Problem. As a team, rate Problem C using the figure below.](

This protocol allowed adequate time for reflection and discussion after students had solved the problems.

**Results from Week 4**

During this focus group interview, students were much less agitated regarding their in-class examinations. Students asked us to switch to sandwiches instead of pizza because of caloric intake.

**Principles** – All students recognized work and energy as concepts covered in these three problems. Five students directly noted the work-kinetic energy equation as the principle involved on their worksheets. All other students used the work-kinetic energy theorem in their solutions. Students explained that their choice of underlying principle is obvious from the example and material covered in class.

Student DM “I know that this equation is used here, and we are given the same information in this problem, just with different numbers and different objects, so it’s the same principle.”

Student AR “It’s also the only part of this chapter we covered in class so far, so you weren’t trying to confuse us with stuff we haven’t covered yet, right? So, yeah, it’s work-kinetic energy.”
**Text Editing** – All of our students identified a distance as being irrelevant for solving their problems, but not all students recognized which of the two distances given was correct. Four students did not choose any of the information as being irrelevant and stated that they were unsure as to what was unnecessary until they knew their solution was correct. Of the six students that chose a distance, only three were sure they chose the correct distance.

Interviewer “So, you have the 0.8 m distance selected as irrelevant. How did you come to the conclusion that this was not necessary?”

Student SS “I have too many distances..haha. I don’t know, I think this one needs to be added to the other one, or maybe this one or the other one is just not needed.

Student MD “I could figure out which one is necessary if you give me the solution, Fran, solution…Fran?”

**Similarities and Differences** – When asked to identify the similarities and differences, most students communicated that the problems A, B, and C were very similar. Students determined that the problems were different because one problem asked them to solve for velocity while the other asked for them to solve for a force. Students also mentioned that the problems were similar because both questions required the use of the same equations and objects in both problems were being stopped by another object (For problem B, this is a true statement, but it is not pertinent to the solution of the problem. We focus on an arrow leaving a bow, not the arrow hitting the target). From Table 5.2 below, there were two students that rated the similarity between problems at or lower than 2.5. These students verbally listed the same similarities and differences as the rest of the focus group learning interview participants, and it was clear from these students’ worksheets that the equations and concepts remained primary similarities, while solving for different values remained primary differences. It is unclear as to why the similarity ratings for these two students were lower. It is also interesting that these same two students, rated the similarity between problems A and B as the lowest measured against other problem comparisons.
Table 5.2 Similarity and usefulness ratings for week 4.

<table>
<thead>
<tr>
<th>Participants</th>
<th>C</th>
<th>F</th>
<th>M</th>
<th>M</th>
<th>S</th>
<th>D</th>
<th>M</th>
<th>A</th>
<th>R</th>
<th>S</th>
<th>A</th>
<th>E</th>
<th>M</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Similarity between A and C</td>
<td>3</td>
<td>3</td>
<td>2.5</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>Similarity between B and C</td>
<td>3</td>
<td></td>
<td>3.5</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Similarity between A and B</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3.5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usefulness rating for C on Problem A</td>
<td>4</td>
<td>4</td>
<td>2.5</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>Usefulness rating for C on Problem B</td>
<td>4</td>
<td></td>
<td>4</td>
<td>2.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Usefulness Rating** - Students were also asked for the first time to rate the example problem on how useful it was to help them solve their unsolved problem A or B. Individual ratings can be seen in Table 5.2 above. Students’ average rating for usefulness is a 3.5 out of 5, where 5 is most useful and 1 is least useful. Students were not asked during this interview to explicitly talk about why they rated the problem as useful or not useful, but a note was made by moderators to ask students about the rating they gave for future focus group learning interviews.

**Problem Posing** - Students worked on the creation of their own problem. Four students became confused between the distance traveled by an object while a force acts on it, and the distance traveled by an object moving freely. For example, the distance a cannon ball travels after leaving the cannon versus the distance the cannon ball traveled through the barrel of the cannon while the powder is ignited.

Three problems are posed by student groups, and all three problems are solvable by other groups. One of these three problems was actually correctly solvable, while the other two required additional information about the distance a force acted on the object in motion. All problems created had a change in features (ie, cannon or a quarterback throwing a football instead of an arrow or baseball). The cannon problem was facially most similar to the arrow being shot, but the problem gave information pertaining to the cannonballs’ trajectory and initial force applied. It did not give any information regarding the distance the force was applied while the cannonball traveled through the cannon. Thus the problem was underspecified.
The quarterback problem was structurally very similar to the arrow problem and solvable. The authors of the quarterback problem also take the time to add irrelevant information to the problem that did not overspecify the problem. The third problem uses a bow shooting an arrow, but the problem asked for velocity instead of average force, and the distance given in the problem was not the distance the force was being applied. This problem was thus underspecified. Figure 5.11 displays the cannonball problem created by two of our focus group participants. White blocks on the figure indicate the elimination of names associated with our participants. The rest of the problems are in Appendix C-2 under the weekly summaries.

Figure 5.11 Example of problem created by students in week 4.
5.3.5 Focus Group Interviews-Week 5

Methodology

There were only two minor differences between the week 4 and week 5 protocol. We continued to ask students to discuss the similarities and differences. Additionally, we asked students to rank the similarities and differences between Problems A, B, and C in order of importance to obtaining a solution. Figure 5.12 shows the problem pair and example problem given for week 5. The topic covered for FOGLI 5 was rotational motion. Figure 5.13 shows the table provided for students to complete the similarities and differences task. For full solution to problem C and the full protocols used for this week, see Appendix C-2. The phrasing of each step was also changed to make the instructions clearer.

Figure 5.12 Constrasting cases presented to students in week 5.

Problem A
A pendulum bob of mass \( m_1 = 0.80 \) kg hangs 0.40 m above the ground. The pendulum bob is attached to a fixed support at a frictionless pivot using a rigid rod of length 1.0 m and negligible mass. The rigid rod is struck by another pendulum bob of mass \( m_2 = 0.50 \) kg that is attached to the same supporting point with a 0.70 m. The second pendulum bob strikes the supporting rod of the first pendulum at a speed of 2.0 m/s and sticks to it. What is the speed of the two pendulums sticking together immediately after the collision?

Step 1: Identify the principles involved in the problem.
Problem B
A flat horizontal disk of mass $m_1 = 4.0 \text{ kg}$ and radius 0.20 m rotates on a frictionless bearing at 30 revolutions per minute. A second disk of mass $m_2 = 0.50 \text{ kg}$ and radius 0.10 m is not rotating. It is released from rest and strikes the first disk at a vertical speed of 2.0 m/s. The second disk sticks to the first disk after falling onto it. Find the angular speed of the two disks sticking together.

Step 1: Identify the principles involved in the problem.

---

Problem C
The figure shows a device used to determine the speed of a steel pellet. The pellet is fired horizontally into a catcher, which is mounted on a disk. The disk is mounted on a frictionless bearing so that it spins freely about its axis when the pellet gets lodged into the catcher. By measuring the number of revolutions per minute of the disk it is possible to calculate the speed of the pellet.

When a 0.01 kg pellet is fired into the catcher, the disk rotates at 6 revolutions per minute. The catcher is at a horizontal distance of 0.08 m from the axis of the disk and a vertical height of 0.12 m above the surface of the disk. The disk has a mass 0.20 kg, radius 0.10 m and thickness 0.03 m. The mass of the catcher and its support structure are negligible. Find the speed of the pellet.
We also asked students to go over the example Problem C individually, rather than have one of the two moderators explain the solution on the chalkboard. After students asked any questions they had pertaining to the solution, we provided them with Problems A and B.

**Results from Week 5**

Students spent the first 15 minutes of the FOGLI to look back and forth between their problem and the example problem. Though students were not looking to their books for more examples, students were taking longer than last time to get through these problems. Two people in separate groups were capable of solving the problems, and tried to help their own partners with solving the problem.

Eventually, everyone completed their solution, and enough time remained for students to discuss the similarities and differences between the three problems A, B, and C. Some students began to get off task as the moderators wandered around asking students to summarize their similarities and differences. Time ended before students
could complete the full set of tasks. The problem which asked students to solve for velocity instead of an angular velocity tended to cause slightly more confusion than the other problem.

**Principles** – Students immediately identified angular momentum as being an integral concept for these problems. When asked to clarify how they knew this to be true, students cited the moment of inertia given in the problem statement as being a “dead give away.” Students also stated that you could tell by looking at the Problem C example, as the conservation of angular momentum is used to solve for the velocity.

  Interviewer “How did you identify your concept involved in this problem?”
  Student MR “Conservation of angular momentum is used in the example?”
  Student MM “Yeah. That’s sort of a dead give away.”

**Text Editing** – Students disregarded the text editing task for this interview while they tried to work the solution to the problem.

  Student MS “I would prefer to do this part later, I’m still trying to figure out how to even go about solving this problem.”

**Similarities and Differences** – The primary similarity identified by students was that all problems incorporated conservation of angular momentum. Problem A and B remained different with respect to the lengths, masses, and moments of inertia. One student also noted verbally that there was another difference with respect to the final motion of the objects. In both problems there was one object hitting another object and changing the second objects’ velocity. The first object continues to have angular momentum in problem A, while the first object in problem B does not have angular momentum.

Students verbally agreed that the example problem was fairly close to problems A and B. Table 5.3 shows the ratings students submitted for the similarities between problems.

**Table 5.3 Similarity and usefulness ratings for week 5.**

<table>
<thead>
<tr>
<th>Participants</th>
<th>C</th>
<th>F</th>
<th>M</th>
<th>M</th>
<th>S</th>
<th>M</th>
<th>D</th>
<th>D</th>
<th>A</th>
<th>R</th>
<th>S</th>
<th>S</th>
<th>A</th>
<th>M</th>
<th>E</th>
<th>MR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Similarity between A and C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.25</td>
</tr>
<tr>
<td>Similarity between B and C</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3.5</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Similarity between A and B</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.17</td>
</tr>
<tr>
<td>Usefulness rating for C on Problem A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.5</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>

136
### Usefulness Rating

Students rated the example problem as being useful for helping them solve their Problem B. Students given problem A rated the example as slightly less useful. When students were asked to clarify why the problem was not as useful for problem A, students stated it was because the moment of inertia for the pendulum problem was different than the example and that added extra steps to the solution.

Student CF “I actually have to do more work…not like physics work, but math work. This sucks. It makes this example less useful to me.”

### Problem Posing

Students did not have enough time during the week 5 focus group learning interview to complete the problem posing task. One pair of students began drafting a problem on their sheet but were unable to collaborate long enough with one another to complete the task. The drafted problem looked facially and structurally similar to problem C, but with an extra disk of radius r is added on top of a rotating disk of radius R. The students were unable to complete the problem, but it looked like disk 1 seemed to serve no purpose other than to rotate frictionlessly with disk two.

### 5.3.6 Focus Group Interviews-Week 6

#### Methodology

The protocol remained the same as week 5 for the remainder of the semester. The problems given for week 6, covering pressure in fluids, are provided in the appendix C-2, and also shown in Figure 5.14.
Results for Week 6

Before beginning their own problems, students were given the opportunity to ask questions about the example problem C. For the most part, up until now, questions had been minimal with respect to the solutions. For this week, seven students all separately asked whether the pressure would be the same at all parts of the bottom of a pool. One of the two moderators answered, “yes, it is only dependent on the depth, or water level.” The same moderator, noticing a trend in the questions asked, stated to all students in the focus group learning interview that the height of the water column is necessary to measure the pressure.

Students began working on problem A and problem B. There were two questions by students given problem A regarding the mechanics of a hydraulic lift. One of the two moderators determined a picture would alleviate some of the confusion, and drew a picture of a car on a hydraulic lift on the chalkboard for everyone to view.
These problems were solved more quickly than the set from the previous week. Everyone was able to create their own problems and hand them to another group to solve. Seven of 10 students also stated during this interview that they would be taking General Physics 2 in the future. General Physics 2 is the second semester introductory algebra-based course. One person said they would be taking it in the summer. Another student said they would be taking it in the following spring and not in the fall. This information, though irrelevant for understanding students problem schema, was important with regards to the next phase of our project. Up to this week, it was unknown as to whether the study would continue with the same set of students into the next sequence. Unfortunately, we were losing too many students from our cohort, and we eventually made the decision to replicate our current study rather than follow this cohort of students into the next semester.

**Principles** – The concept was easily identified by everyone as pressure. However, more time was spent by four students to explain that their problem only involves specific formulas related to pressure, and not as many as the example problem.

Student MM “Well its obviously pressure, but they are all pressure. The difference is that this one is pressure using $\rho gh$."

**Text Editing** – Students were capable of selecting the irrelevant information given in the problem without any hesitation. When asked how they were able to come up with the irrelevant information so quickly, students replied that they could figure it out by what was not given in problem C. Moderators quickly made a note not to include irrelevant information that could be so easily picked out through simple comparison with problem C.

**Similarities and Differences** – Students identified the following similarities and differences: The formulas used were the same between problem A and C, and B and C. Both questions were physics questions. The problems A and B were pieces of example Problem C. The tube for the hydraulic lift was analogous to the tube coming out of the barrel in problem C, and the pool was analogous to the barrel in problem C. Thus, students felt problem C was partly similar to A and B, but A and B were not necessarily similar to one another. The worksheets and ratings provided evidence of this conclusion as the similarity between problems A and B were only similar in terms of both being
problems involving pressure, whereas, A and C, or B and C, were more similar by the number of equations that remained similar between them.

Student AR “Problem C was exactly like problems A and B, so it was like plug and chug. You just need to know which parts of the problem C go with which parts of these problems [problems A and B], but that’s easy to do.”

Interviewer “How did you determine what parts were necessary for your problem?”

Student AR “You can just tell. My problem has a pool and I need to know the pressure at some level in the pool. That’s like finding the pressure at the top of the barrel. This part [of problem C] uses the equation I need.”

**Table 5.4 Similarity and usefulness ratings for week 6.**

| Participants | C  | F  | M  | M  | S  | D  | D  | A  | R  | S  | A  | E  | M  | R  |
|--------------|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Similarity between A and C | 1.5 |  | 3  | 1  | 2  |  |  |  |  |  |  |  |  |
| Similarity between B and C | 4  | 2  | 3  | 2.5 |  |  |  |  |  |  |  |  |  |
| Similarity between A and B | 1  | 1  | 2  | 2  | 2  | 1  | 2  | 2  | 0.5 | 1  |  |  |  |
| Usefulness rating for C on Problem A | 3  |  | 2.5 |  | 3  | 1  |  |  |  |  |  |  |  |
| Usefulness rating for C on Problem B | 4  | 3  | 4  | 4  |  |  |  |  |  |  |  |  |  |

**Usefulness Rating** – Problem C was rated most useful for problem B, with an average rating of 3.5. Problem C was given a mean usefulness rating of 2.3 by students solving problem A. During conversation with students, it is unclear as to why problem C was considered less useful for problem A than problem B. Students continued to state they used problem C to help them solve the problem, but that there remained significant differences between the two problems.

Student DM “Problem A has a pool of water and problem C has a pool of water with a piston which is totally different.”

Student EJ “yeah, there are different equations included with problem C. It’s more complicated.”

**Problem Posing** – Two problems were posed by student groups for week 6. A whale is lifted out of water on a hydraulic lift in the first problem (See Figure 5.15). It was facially and structurally most similar to Problem A, requiring the same mathematical
procedure to solve and the hydraulic lift was still present. Dimensions of the pool were provided to the solvers, but there was insufficient information to complete this problem. The student solvers came to the same conclusion.

Figure 5.15 Example of problem posed in week 6.

The swimming pool problem posed was a variation of Problem B and C, as it included a cylindrical tank with a tube sticking out like problem C, but the problem procedure was similar to B. The swimming pool problem also interchanges information given and information asked for. This problem posed was solvable and the student solvers assigned to this problem were capable of completing the problem. All problems posed for all weeks are found in Appendix C-2.
5.3.7 Focus Group Interviews-Week 7

Methodology

The protocol for the focus group learning interview in week 7 remained the same as in week 5 and week 6. The problems given for week 7 covered simple harmonic motion and are provided in the appendix C-2, and also shown in Figure 5.16. It was more difficult to create two problems for this week that shared little surface feature similarities with simple harmonic motion problems. We were constrained to problems only involving spring systems as this was all that was covered during class. In order to create problems that were most surface feature dissimilar, we changed the axis of orientation, generating problems where the springs were fixed to the ground, a sidewall, or the top of an incline. These features carry significant differences in associations with students as inclines are often associated with angles, and the objects in vertical motion are often associated with gravitational effects. But, these particular problems did not require significant changes to their solution with regards to these different features. These problems were not as surface feature different as previous problems, but they remained different enough for the purpose of this interview.

Figure 5.16 Contrasting cases provided for week 7.

Problem A

A 0.20 kg mass is attached to the end of a horizontal spring bolted against a wall. When the spring is fully compressed, the distance from the wall to the midpoint of the mass is 2.1 m. Assuming there is no friction between the mass and the floor, the mass is released and the system begins to oscillate. The equation that describes the motion of this system as a function of time is \[ x = 0.7 \cos(2\pi(3)t) \]

(a) What is the frequency of the motion?
(b) What is the mass of the block?
(c) What is the maximum velocity obtained by the mass?
(d) What is the total energy?
**Results for Week 7**

The significant differences seen between week 6 and week 7 were the difficulty levels of problems provided. The problems for week 5 and 6 were considerably more difficult for students than its successors (problems for weeks 7 and 8) with regards to the complexity, or mathematical work necessary for obtaining a solution.

Students had difficulty focusing for this interview. Since many of our cohort shared majors that require similar electives, many of them were required to dissect an unborn chick from an egg earlier in the day. Moderators asked students to politely stop talking about the dissection. Prior to beginning their worksheets, three students asked the same question regarding the example problem C. These three students wanted to know how a person should decide whether the equation of motion should have a sine or a cosine function?

One of the moderators set up two scenarios involving springs with differing initial starting conditions. The moderator then showed students that the initial condition of the
spring determined which of the functions were selected for the equation of motion. The moderator also made the caveat that sine and cosine functions could be shifted to look like one another, and so it is possible to represent any of their simple harmonic oscillators with a sine or cosine function.

Problems A and B were handed out to students shortly after one of the two moderators announced an error on the first page of problem B. The second question was suppose to read, “What is the amplitude?” Instead it asks for a quantity already given in the problem statement.

**Principles** – Students identified the principle as either simple harmonic motion, or the oscillation of an object. When asked to clarify what they meant by simple harmonic motion, students clarified similarly,

Student SS “….like something moving back and forth. It doesn’t have anything working against it, that’s the simple part, you know?”

Only one of our participants is unable to determine what it means to have simple harmonic motion.

Student MR “It’s when you have an object that oscillates. You know, like a spring.”

Interviewer “How is simple harmonic motion different from general oscillatory motion?”

Student MR “It uses simple objects, like a spring or a pendulum.”

**Text Editing** – Students are also capable of determining the irrelevant information given in the problems. For this material, students have a lot of difficulty describing how they knew what information was irrelevant. They did not have to solve the problem to find it, but they could not describe to me why they knew it was unessential.

Student MD “I don’t know, Fran! Sometimes you just know.”

Student EJ “I think it’s because I did enough homework problems.”

Student MM “Umm. Because…it doesn’t tell us anything useful. [prompted to continue with this thought] I don’t know, it would be useful if you were calculating potential energy, but that’s not really needed here because you already know enough to find the equation of motion.”
Similarities and Differences – Students found problem B to be very similar to problem C because it uses the same formulas (one person said it was a simple plug and chug). Most students agreed problem A was less similar to problem B, but still similar.

Student AM “Problem A is slightly more difficult because it is not identical to problem C, but still not difficult.”

When asked to clarify what they meant by “less similar”, student AM stated that “Problem A was given the equation of motion while Problem C had asked for it.”

Students rated problem A, B, and C all similar to one another, though problem B was more similar to problem C because the procedure for solving problem B remained exactly the same as problem C.

Table 5.5 Similarity and usability ratings for week 7.

<table>
<thead>
<tr>
<th>Participants</th>
<th>C</th>
<th>F</th>
<th>M</th>
<th>M</th>
<th>S</th>
<th>M</th>
<th>D</th>
<th>D</th>
<th>A</th>
<th>R</th>
<th>S</th>
<th>S</th>
<th>A</th>
<th>M</th>
<th>E</th>
<th>J</th>
<th>M</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Similarity between A and C</td>
<td>3</td>
<td>3</td>
<td>3.5</td>
<td></td>
<td></td>
<td>3</td>
<td>3</td>
<td>2.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Similarity between B and C</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3.5</td>
<td>3</td>
<td>3.5</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>Similarity between A and B</td>
<td>2.5</td>
<td>3.5</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usefulness rating for C on Problem A</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</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>
<td></td>
<td></td>
</tr>
<tr>
<td>Usefulness rating for C on Problem B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</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>
<td></td>
</tr>
</tbody>
</table>

Usefulness Rating – Students rated problem C useful for both problem A and problem B. Problem B rated higher overall because the problems asked for the same quantities, where problem A interchanged some of the quantities or equations given that were asked for in problem C. Overall Problem C was given a usefulness rating of 4.0 for problem B, and a usefulness rating of 3.0 for problem A.

Problem Posing - Students were able to create their own problems, but they were not discussed as a general group because we ran out of time. There were 5 problems posed for week 7. In all five problems posed, the students varied the direction under which the spring would move like problems A, B, and C, so in other words, the springs may be fixed horizontally, vertically, or at an angle. The problems asked for the same variables as those asked for in previous problems A, B, and C. In other words, students were capable of creating multiple combinations of spring direction and statement selection. All problems contained sufficient information to solve the problems, and all student
groups assigned to solve these problems posed were capable of solving the problems. Two of the problems posed remained unsolved only because they were completed at the very end of the focus group learning interview and other groups were not given enough time to solve the problems. It is interesting to note that the facial features associated with these problems did not vary at all since all the systems among the posed problems remained blocks and springs.

**Figure 5.17 Example of problem posed during week 7.**

![Image of a hand-drawn diagram]

It makes sense that the students kept facial features similar or the same as compared with Problems A, B, and C because it best matched the instructions given to them. However, if we compare with the previous week, there were significantly more surface feature changes including the addition of a whale and people.

### 5.3.8 Focus Group Interviews-Week 8

**Methodology**

Week 8 was the final focus group learning interview for the semester, and material covered included standing waves and resonance. The problems provided to students for
week 8 are shown in Figure 5.18. For full protocols and solutions, please see Appendix C-2.

Figure 5.18 Contrasting cases provided for week 8.

<table>
<thead>
<tr>
<th>Problem A</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 2.0 m long open-ended organ pipe when filled with an unknown gas at 20°C resonates at two successive frequencies of 400 Hz and 500 Hz.</td>
</tr>
<tr>
<td>(a) What is the fundamental frequency?</td>
</tr>
<tr>
<td>(b) What is the velocity of sound in the gas at 20°C?</td>
</tr>
<tr>
<td>(c) What is the number of the harmonic corresponding 400 Hz and 500 Hz respectively?</td>
</tr>
<tr>
<td>(d) Sketch the wave for the harmonic at 400 Hz and 500 Hz.</td>
</tr>
<tr>
<td>(e) What is the fundamental frequency and the next three higher harmonics if you now close the open end of the pipe?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Problem B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A close-ended organ pipe when held in air at 20°C resonates at two successive frequencies of 375 Hz and 525 Hz.</td>
</tr>
<tr>
<td>(a) What is the fundamental frequency?</td>
</tr>
<tr>
<td>(b) What is the length of the pipe?</td>
</tr>
<tr>
<td>(c) What is the number of the harmonic corresponding 375 Hz and 525 Hz respectively?</td>
</tr>
<tr>
<td>(d) Sketch the wave for the harmonic at 375 Hz and 525 Hz.</td>
</tr>
<tr>
<td>(e) What is the fundamental frequency and the next three higher harmonics if you now open the close end of the pipe?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Problem C</th>
</tr>
</thead>
<tbody>
<tr>
<td>A pipe is designed to produce successive harmonics at 100Hz, 150Hz, 200Hz in air at 20°C.</td>
</tr>
<tr>
<td>(a) What is the length of the pipe?</td>
</tr>
<tr>
<td>(b) Is the pipe open or close ended?</td>
</tr>
<tr>
<td>(c) What is the fundamental frequency?</td>
</tr>
<tr>
<td>(d) What is the number of the harmonic corresponding to 100Hz and 150Hz respectively?</td>
</tr>
<tr>
<td>(e) Sketch the wave pattern for the harmonic at 100Hz and 150Hz respectively.</td>
</tr>
</tbody>
</table>

Results for Week 8

Five students did not attend part or all of the final interview. In an attempt to keep students’ focus, one of the moderators addresses a question about the open and closed tube diagrams. During this explanation, it became apparent to that moderator that there is an error in the solutions normally handed out at the very end. The two moderators fixed the error by hand while students worked on their unsolved problems.
Students were taught a math trick in class to do these type of problems. The trick made the problems given for this week trivial, and also rendered problem C useless.

**Principles** – Students recognized that the problems solved for fundamental frequencies and were open and closed harmonic waves in pipes. Most students did not share the same ‘text-book’ name for the principle, but instead described the system.

Student AR “This is just a closed pipe with harmonic wave motion so the wave is moving back and forth and I can describe its nodes.”

Student EJ “Hey that’s a good explanation. Me too. Except mine is open ended and I have to take into account the gas that’s in the pipe because mine has some unknown gas.”

Interviewer “How does the gas have an effect on your solution?”

Student EJ “Uhh...ff...so its how the wave moves..it needs gas.”

**Text Editing** – For the final interview, students did not answer the text editing question on their worksheets. When prompted to answer the question verbally, three students given problem B correctly stated that the problem did not contain unnecessary information. Three of four students given problem A correctly determined that the temperature of the unknown gas is unnecessary. Only one student actually wrote it down on their worksheet.

Student EJ “The temperature of the unknown gas isn’t really needed because we can calculate the velocity anyways.”

The student that did not believe there was irrelevant information given in problem A eventually determined they were incorrect after listening to one of the others in the group explain their problem solution.

**Similarities and Differences** – Students discussed the similarities and differences between problems A and B. Problems A and B were cited as similar because they all solved for frequencies and the formulas were similar. Problems A and B were different because A and B had different gases inside the pipe and one was open ended while the other was closed.

Student CF “Problem A is similar to problem b, and problem b is similar to problem c, and so on. They are all the same problem except the number in front changes depending on if you have open or closed pipes.”
showed us a really easy way to do these problems…and I don’t really have to look at problem C …just to pull equations off of it.”

Problem C was determined to be equally as similar to problem A or B, as problem A and B were similar to one another.

Student AM “Problem C uses the same formulas as B and A. All the problems are a little different from one another, but they are all pretty close too”

[Others in the room nodded in agreement as Student AM made her statement aloud]

**Table 5.6 Similarity and usability ratings for the final week 8.**

<table>
<thead>
<tr>
<th>Participants</th>
<th>C</th>
<th>M</th>
<th>M</th>
<th>S</th>
<th>M</th>
<th>D</th>
<th>M</th>
<th>A</th>
<th>R</th>
<th>S</th>
<th>A</th>
<th>E</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Similarity between A and C</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Similarity between B and C</td>
<td>3</td>
<td></td>
<td></td>
<td>2.75</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Similarity between A and B</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2.5</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usefulness rating for C on Problem A</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Usefulness rating for C on Problem B</td>
<td>1</td>
<td></td>
<td></td>
<td>2.5</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Usefulness Rating** - When asked what resource was most useful, students felt the final examination equation sheet and problem C were both useful for Problem A, while the equation sheet was more useful for Problem B. Students made it clear that problem C was only good for the equations presented on the sheet, as they had a better way of solving the problem. The actual ratings for problem C’s usefulness ranged depending on whether students found that the equation sheet was more useful or problem C’s equations were more useful. The usefulness ratings for problem C can be found in Table 5.6.

**Problem Posing** - Students designed three problems that require the solver to figure out whether a pipe with given frequencies was open or closed. One group tried to make their problem unique by adding monkeys (See Figure 5.19).
Again students created problems that shared several deep structure and surface feature similarities with problem A, B, and C. All problems created for this week were solvable and the solutions were completed by all student solvers.

Finally, students were asked how they felt about being given a problem C for the last five weeks.

Student AM “It’s a lot easier to see an example worked out before attempting a similar problem.”

All students agreed that textbook examples and examples from lecture were useful for solving homework problems.

5.3.9 Summary of Results

As described earlier, our protocol for the group learning interviews did not stabilize until week 4. Over the first three weeks we changed the protocol significantly toward providing increased procedural scaffolding to relieve students’ cognitive resources to focus on reflecting about the problems rather than on simply solving them.
In the last five weeks (weeks 4 through 8) of the group learning interviews we were able to provide sufficient scaffolding that would allow time for reflection. The data collected from students’ worksheets, field notes collected from the moderators, and problems posed by students each week can be found in Appendix C-2.

**Principles**

Students were asked to write out the principles involved in the problem they were given to solve, and also to determine whether there was missing, irrelevant, or sufficient information not given/given in the problem statement. It was extremely important to pose students specific questions asking them to enunciate principles of a problem and provide them with a concrete structure to facilitate reflections on similarities and differences between problems. Until the end of the focus group learning interviews, students continued to have difficulty enunciating principles of a problem past the chapter titles. Verbal confirmation of a principle between moderators and the participant was necessary and students were often asked to first describe the event which takes place in the problem and then explain how that event maps onto the principle they originally selected. The process of this verbal confirmation was very quick, but necessary. If one were to study the written materials alone, it would be unknown as to whether students were capable of recognizing the principles involved or if they were just choosing the principle based upon the principles involved in the example knowing that they would be the same or similar.

   Interviewer “What were the principles involved?”
   Student “Harmonic wave motion.”
   Interviewer “Could you explain how harmonic wave motion best describes this problem’s principle?”
   Student “You have a wave travelling through the pipe and it goes back and forth.”

**Text Editing**

The text editing task required students to determine whether there was any information given or not given that was irrelevant or missing in the problems,. It is important to note, that during one of the focus group learning interviews, students were
capable of selecting given irrelevant information without having to solve the problem. It was possible for students to extrapolate irrelevant information given in an unsolved problem statement by looking at the matching up individual given quantities with the solution of the example problem. The moderators picked up on this extrapolation and made corrections for it in all subsequent interviews. From that point forward, more effort was put into problem creation such that the examples would give away less information regarding irrelevant and missing information. This was done by selecting irrelevant information that would match the type of information necessary for the problem solution.

For example, if a displacement was a necessary physical variable in the problem, an extra distance quantity might be given such that the students could not rely on units alone for extrapolating the unnecessary information.

Students would often leave this section blank initially and go forward to solve the problem. Once a solution was obtained, students would go back to fill out the text editing section. Moderators would ask students to fill out the section before solving the problems, but it was difficult to force students to do so. Finally, an agreement was made that students would hypothesize as to what information may be irrelevant and they could go back to change their hypothesis if they were incorrect. It was during verbal discussion with the moderators that, for the most part, students were capable of selecting the irrelevant information given in the problem, but it was often difficult for students to communicate why it was deemed unnecessary.

Student SS “I don’t know. It just is…like you would need to have one of these information and this one is not relevant.”

The problems given never had missing information, so though we still asked students what information is missing if any, students would either leave it blank or would fill in information that was not explicitly stated, but implied, such as gravitational acceleration.

**Similarities and Differences**

Students were asked to reflect on and then describe the similarities and differences between the two problems A and B each given to one of the students in each pair. They were also asked to compare each of these problems with problem C, the solved example. While comparing problems, students often recognized the commonality
of the underlying principle. The similarities cited by students were based on the deep structure of the problems. Although students also pointed out similarities in surface features, wherever applicable, they often ranked these similarities as being less important than the similarities in the deep structure.

Student CF “They’re both problems using the same equations. They look different, but they solve out the same way. It’s just Work.”

The differences between problems identified by students for the most part focused on the surface features.

Student DM “You have to convert this from angular velocity to linear tangential velocity, so that’s different.”

Some of the differences might have gone beyond just surface features and might have affected the underlying mechanism of solving the problem, though students pointed out these differences in their comparison as well.

Student MM “This one uses the moment of inertia of a disk with radius r, and this one uses the moment of inertia for a sphere. It’s just a factor of ½ mathematically, but you can’t forget about that.”

The similarities and differences that students focused on seem to remain consistent for the four interviews following the finalized protocol.

**Usefulness Rating**

In addition to comparing various problem pairs, students were also asked to rate the usefulness of the solved example – problem C in helping them solve problems A and B. On average, the usefulness rating of an example was a 3.0, however it varied between 2.5 and 4.0 regularly. Often an example would rate higher for either problem A or problem B than its counterpart, though the rating did not differ between A and B by more than one point on average.

In rating the usefulness of a solved example in solving the problem, students were more likely to find the example useful if the steps in the solution of the solved example mapped directly onto the steps in the solution of the unsolved problem. In other words, students looked for procedural elements in the solved problem and not necessarily elements in the underlying conceptual schema to facilitate their solution of the unsolved
problem. This was interesting to see in week 8, where students felt the example problem C was less useful because the procedure was the full detailed solution and not the shortcut method they learned in class. The dislike toward the longer mathematical procedure is not evidence in itself, but the lower ranking in usefulness shows that students are less apt to study the full solution over a similar procedure which incorporated a mathematical trick. The ‘shortcut’ method was presented by the instructor in class and while it provided an efficient method for solving the problem, it did not help students think about the problem conceptually. The bottom line is that from the students’ perspective, the focus on problem solving continues to be on procedural case reuse rather than schema abstraction.

We learned that it is extremely important to design the solved example appropriately to optimize its usefulness in problem solving. The use of mathematical derivation in a solved example could reduce its perceived usefulness in facilitating the solution of the unsolved problems if the derivation is perceived to be unnecessarily long in week 8. For these instances, it is important to make adjustment on the amount of mathematics that is described in an example procedure such that students are capable of following along without having to make large leaps of faith or skim large sections that might be considered unnecessary. It is an important skill to learn how to extract information from a given problem example when it is not carefully arranged to fit a previously seen solved example, but for this project, our focus was to ultimately determine how novices select their cases to be reused and to assess strategies that might facilitate emphasis on deep-structure characteristics during case reuse. The difficulty level of unsolved problems must be also be carefully adjusted. If the problems are too difficult (such as in week 4) students tend focus on resolution of a challenging problem, not on reflection. If the problems are too easy (such as in week 8), students do not need to reflect on what they have learned from the solved example and how it might be applicable.

**Problem Posing**

After students were prompted to be more careful about under-specifying their problems, it seemed that students were more effectively posing problems that were
Groups also seemed more attentive to the feedback asserted by student solvers. Feedback included things such as spelling errors, poor handwriting, improper grammar, and underspecified problem statements.

It was difficult to surmise what prompted students to add their own original surface features and irrelevant information, but it seemed related to the degree of variance between problems A, B, and C. If problems A, B, and C were very closely surface-feature similar, such as in week 7 and week 8, the problems posed were less creative. If problems A and B were less similar to one another facially, then students brought in more additional features like in week 4 and week 6.

5.4 Individual Interviews

We conducted two individual interviews with all of the students in our focus group learning interviews. As shown in Figure 5.1, the first individual interview was conducted after week 4 of the focus group learning interview and the second individual interview was conducted at the end of the semester after all of the focus group learning interviews had been completed.

The purpose of these individual interviews was to assess the extent to which students’ conceptual schema with regard to problem solving had evolved due to their participation in the focus group learning interviews. We developed problem sets for the individual interviews that contained problems with varying degrees of deep-structure and surface-feature similarities. All problems shared an overarching concept (i.e., conservation of energy). Four distinct problem scenarios were chosen, (i.e., a man pushing a piano, a rock interacting with a spring, etc) each incorporating one of two primary conditions (i.e. no elastic potential energy or elastic potential energy, swinging pendulum motion or oscillating spring) governed by the overarching concept.

Each distinct problem scenario was used to create a set of three problems of varying deep-structure similarities. The first and second problems would be solvable using the same equations, but with different given and unknown information provided in the problem statements. The third problem was conceptually similar to the first two problems but required a different physical principle (work-kinetic energy, conservation of energy) than the first and second. The third problem had surface features similar to the
first and second problem, and the same primary concept remained, but the different principle would critically alter any equation(s) used for problem 1 and 2, making the third problem structurally different compared with problem 1 and 2.

Figure 5.20 Diagram representing two scenario sets.

Because of this creation of problem sets, there always existed two separate sets of problems that could be paired in such a way as to create all variances in surface feature similar/different and deep-structure similar/different. See Appendix C-3 for the full problem sets given in interview 1, and Appendix C-4 for the full problem sets given in interview 2. Figure 5.21 represents how different pair types were formed using the problems represented in Figure 5.20. More specific examples of pairings will be described later in this chapter when we discuss the similarity rating task.
The individual interviews required students to perform three separate tasks all involving the problems described above. Problems used for the ‘identification task’ and the ‘usability task’ were not the same problem set as used in the ‘similarity ratings task’. Students participants were randomly split into two groups such that problems given for tasks identification and usability task for group 1 were the same problems used for the similarity ratings task for group 2 and vice versa. Due to time constraints, the identification task was not completed sufficiently. The results and methodology for the task are described below.
5.4.1 Task 1 - Identification of principles, equations, and irrelevant information

Students were assigned six problems made up of two problem scenarios that did not have corresponding primary conditions. Half of the students were assigned problems that comprised of one set of two problem scenarios and the other half were assigned problems from a different pair of problem scenarios. Problem 1 and 2 of one scenario (conservation of energy) were given first, then problem 1 and 2 of the second problem scenario (conservation of energy including a spring) followed. The last two problems were the third problem of each of the two problem scenarios (Work and energy). Figure 5.22 represents how a given problem set was formed for the identification task. Again, there were four separate scenario problem sets. The example below explains how a given problem sequence was formed using two scenario problem sets.

Figure 5.22 Example of a problem sequence for the identification task.

See Appendix C-3 and C-4 for the full problem sets given for the identification task of interview 1 and interview 2, respectively. Students were asked to identify the principles involved in the problem as described in their own words. They were then asked to identify the equations applicable to the problem solution given the examination equation sheet, and they were asked to identify whether irrelevant information was given in the problem statement and why that information might be irrelevant. Students are also asked specifically not to calculate the solution.
5.4.2 Identification Task – Results

Students were unable to complete this task in individual interview 1 and interview 2 due to insufficient time. It was determined through the first four individual interviews that the time was running over by an average of 20 minutes due to the depth of questioning necessary to provide sufficient explanation of student reasoning on a given task. It was thus decided that the identification task was the most expendable as the task was similar to the first two steps asked to be performed during the focus group learning interviews. The students were asked during the focus group learning interviews to both describe the concepts and principles associated with the problem given to them (step 1), and they were asked to determine whether there was missing, irrelevant, or sufficient information given/not given in the statement (step 2). The first task was an exercise that was designed to assess whether students were capable of deconstructing a problem statement without first solving the problem. It was unlikely that we would gain explicit insights regarding students’ organization of knowledge through this task without an acceptable amount of time for questioning. This interview task was also the only one of the three that resembled tasks performed during the focus group learning interviews. For the remainder of individual interviews, students were instructed to work through the identification task, but were told they may opt out of answering the question regarding the irrelevant information if they were unsure, and no time remained for the interviewer to ask questions regarding answers given. Students’ answers to the first task are described below.

For the first question of the identification task, students were asked to identify the principles that are apparent in a given problem. The concepts and principles given by students are listed below, with their explanation if so given.

**Interview 1**

For individual interview 1, all 10 interviewees were able to determine that the primary concept underlying all the problems was energy. Some were capable of determining that energy equations would be the only way to solve these problems starting from first glance of the first problem. Others needed to look at several problems before
recognizing that energy, not Newton’s second law, was needed for solving these problems.

Three of the 10 students were able to determine the first problem was a conservation of energy problem. Four of the 10 were unsure of whether the first problem was a Newton’s second law problem or a conservation of energy problem. These four each decided they would remain unsure until they could solve the problem themselves. The three remaining students determined the first problem was definitely a Newton’s second law problem by the way it ‘looked’.

Student SS1: “Newton’s second law”
Interviewer: “What was it about the problem statement that tells you the principle is Newton’s second law?”
Student SS1: “I’ve seen this kind of problem before. It was a second law problem.”
Interviewer: “You’ve seen this kind of problem before?”
Student SS1: “Mhmm…not this exact one…it didn’t have a picture but it was the same… piano going up a board.”

Equations selected by students corresponded to the principles and concepts selected with many students referring to equations given on the equation sheet by the section heading they fell under. Since the equations sheet was organized by chapter titles, it was apparent that students would use the chapter headings to select the equations necessary. Students were also careful to discard equations that fell under a chapter category if they felt they were not appropriate for the given situation. Examples for the first problem include centripetal force equations and the work-kinetic energy theorem. Of the four interviews where there was time to discuss answers with students, the first student interviewed responded to an interviewers probe as follows:

Interviewer “Could you explain a little more as to why you chose the principle as the work-kinetic energy theorem?”
Student “Yah. There are givens and unknowns, and the givens and unknowns are all numbers that are used for work problems.”
Interviewer “Ok. How might you…”
Student “Oh! wait! no…oh no nevermind.”
Interviewer “What? hey, explain to me what you were just thinking there!”

Student “Haha..right! you physics people are always trying to get in my mind.
[wiggles finger at interviewer] I just thought this might be a different kind of problem…[flips book page to work-kinetic energy section of Giancoli] …but then I reconsidered my first argument. [continues to flip through book]”

Interviewer “Your first argument?”

Student “That all my numbers fit this equation. Well, they don’t…but a lot of them fit into this equation. This one won’t work [conservation of energy formula with spring constant] because it has a spring in it. Am I right? [Looks at interviewer for a few seconds]”

Interviewer “Whether your right or wrong doesn’t matter…I want to hear your reasoning.”

Student “I don’t know, I would want to solve it to figure out what I actually need. Can I solve it before moving on?”

Interviewer “Nope!...”

Conversation with the next three students, and yielded shorter but similar statements made by the rest of our participants, it was clear that equation plug and chug method for selecting a ‘principle involved’ was commonly used by these students.

By the second problem, which resembled the first problem in all but the quantity being solved for, all four students that were initially unsure about the concepts and principles covered were able to identify the concept as conservation of energy. All four students cited the information given in the statement and was stated that it was sufficient to use conservation of energy which as the easier calculation would be preferable. Two of these four students also determined that it might be possible to obtain a solution using a different method, like Newton’s second law, but it would not be necessary. The original three students appeared to be convinced that forces needed to be calculated for first problem. They concluded that Newtons’ 2nd law remained appropriate for second problem. The three students that determined conservation of energy was appropriate for problem 1 also concluded that conservation of energy was appropriate for second problem.
By the third problem in the problem sets, a problem that was deep-structure similar to the first two, one of the three students that were originally convinced that forces needed to be calculated, recognized that the spring force could not be calculated and thus determined that the problem might be solvable using energy equations. One student that determined conservation of energy was necessary for the first and second problem, determined Newton’s second law was necessary for the third problem because a spring was introduced and therefore Hooke’s law must be applicable.

By problem 4, all students determined that energy was a key concept involved in the problems given, but not all were sure that it was the only concept apparent. Five of the ten students stated that the problem would use conservation of energy and would require forces because a spring has been introduced.

Students continued to cite energy as a key concept for problems 5 and 6, but none stated that forces are required. Particularly for the last two problems, students cited the equations as direct proof as to whether a concept was necessary. If the information given in the statement could be plugged into an equation, and another quantity may be obtained that is possibly useful for gaining the solution, they determined it as a ‘worthwhile’ equation.

It is important to note that during this interview, students would often make caveat statements regarding the principles and equations, forewarning the interviewer that the principles and equations were being chosen for their promise and that they are quite possibly incorrect. All students recognized that problems 5 and 6 were more ‘difficult’ than previous problems, and that there were ‘different steps’ involved in the problem solving process. It was difficult for the students to commit to choosing equations while asking them not to solve the problem.

**Interview 2**

For the second interview, students were asked to repeat this task, but with a new sets of problems. Simple harmonic motion was selected as the underlying concept manifested in all of the problems used in interview 2.

There were distinct differences in how students approached their answer for this task in interview 2. The concept was identified immediately by all students as simple
harmonic motion. Furthermore, all students expressed this concept as the ‘principle involved’. Unlike interview 1, students spent time to describe the event which occurred in the problem and noted the motion that may be described as a simple harmonic oscillator.

“It’s a fixed spring that’s moving along the incline. I don’t have to worry about friction, and let’s see….I don’t have to worry about the angle because I am given enough info about the motion..it’s vibrating 3 times per second…I know it’s to be S.H.M [simple harmonic motion].”

The above comment was not prompted by the interviewer, but the comment was not surprising as a similar task was asked of students for the focus group learning interviews. Even more interestingly, four of 10 students spent extra time selecting equations that could be used for a solution in the order necessary for the equations to be used to obtain a solution. Six of ten students (including the four above) also spent time determining how the formulas may be refined to include initial conditions given in the problem statement. Thus, the equations selected from the sheet had either some of the quantities plugged in, or the trigonometric functions were selected to fit the initial starting position of a given harmonic oscillator.

Although this attention to initial conditions is promising, it is important to note that this may simply be an effect seen for simple harmonic oscillators. The concepts and principles involved in interview 1 were quite different from interview 2. It would be interesting to see whether this effect would occur again given different concepts and principles for the second interview.

For a future study of students’ selection of principles and equations, it would be important to ask more questions regarding the uncertainty in their decisions. There is a lack of confidence in their own ability to definitively select a principle or set of principles that best exemplify the event which occurs in the physical system described by the problem statement. These observations are interesting because it is apparent that there is a disconnect between the ‘physical observation’ and its connection to the ‘principle’ involved. None of the 10 students in this study use language that would allow us to infer any deeper meaning associated with their selection of a principle other than that the quantities given in a problem statement fit or do not fit the equation or equations that
describe the principle mathematically. For example, students that mentioned problems in
week 6 were fluid problems cited the pressure values given in the problem statements.
Individual problems are discussed through differences between fluid equations used, not
between the scenarios themselves. Though this disconnect is less apparent in interview 2,
it is difficult to determine whether the treatment is affecting students’
communication/reasoning or if it is simply an affect resulting from students’ perceptions
of harmonic oscillators.

5.4.3 Similarity Ratings Task - Methodology

During the individual interviews students were asked to rate the similarities
between contrasting problems of varying deep structure and non-deep structure
similarities. Research by Chi (Chi, Feltovich et al., 1981) has shown that students tend to
group problems based on surface features, while experts group problems based on their
deep structure. Similarly, Hardiman (Hardiman, Dufresne et al., 1989) showed that
surface similarities between problems could interfere with experts’ classification of the
problems. Our tasks were different from those presented by Chi in her research. Rather
than ask students to categorize the problems we presented students with pairs of problems
and asked them to rate the similarity of each pair on a five-point Likert scale. Each
student was presented with eight pairs of problems. The problem pairs of were
constructed from problems that had facial similarities/differences and principle
similarities/differences. The term facial similarity/difference corresponds to surface
similarity/difference, while the term principle similarity/difference corresponds to deep
structure similarity/difference.

All four combinations of facial/principle similarities/differences were created.
These are labeled problem pair types A, B, C, and D as defined in the Table 5.7 below.

Table 5.7 Problem pairs for the similarity rating task.

<table>
<thead>
<tr>
<th></th>
<th>Facial Similarity (FS)</th>
<th>Facial Difference (FD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principle Similarity (PS)</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Principle Difference (PD)</td>
<td>C</td>
<td>D</td>
</tr>
</tbody>
</table>
Students were presented the problem pairs in order A, A, B, B, C, C, and D, D. Students were not allowed to backtrack and change their similarity rating for any pair until the end of the sequence when they were given the opportunity to review their ratings for all pairs and decide whether they wanted to revise any of the similarity ratings. Figure 5.23 through Figure 5.26 below show examples of the similarity rating tasks used in the study in Interview 1.

**Figure 5.23 Type A problem pair: Facial similarity [FS] (both roller coasters) and principle similarity [PS] (both are conservative systems).**

![Type A problem pair](image)

**Figure 5.24 Type B problem pair: Facial difference [FD] (roller coaster vs. gun) and principle similarity [PS] (both are conservative systems).**

![Type B problem pair](image)

**Figure 5.25 Type C problem pair: Facial similarity [FD] (both roller coaster) and principle difference [PD] (conservative vs. non-conservative).**

![Type C problem pair](image)
5.4.4 Results – Similarity Ratings

Students who participated in the focus group learning interviews were also interviewed individually twice in the semester – first after completing the focus group learning interview in week 5 and the second time was toward the end of the semester after completing all of the eight focus group learning interviews.

We focus on students’ ratings in the problem similarity tasks on the four types of problem pairs of type A, B, C, and D. Recall that the problem pairs contained problems that shared either or both facial similarities/differences or principle similarities/differences as described in Table 5.7. Examples of problem pairs are shown in Figures 5.23 through 5.26.

Before describing the students’ ratings, we describe how we believe an ‘ideal expert’ would rate these problem similarities. Our hypothetical ‘ideal expert’ should focus exclusively on the principle similarities/differences and not at all on the facial similarities/differences between problems. Thus, this ‘ideal expert’ should rate problem pairs A (facial similarity, principle similarity) as well as pairs A (facial difference, principle similarity) as equally high on the Likert scale. This is because our ‘ideal expert’ is completely sensitive to the similarities/differences in principle and completely blind to facial similarities/differences, so although the problems in pair B are facially different, this hypothetical ‘ideal expert’ would rate the problems as being almost as similar as the problems in pair A. Based on the same reasoning, this hypothetical ‘ideal expert’ would rate pairs of type C and D equally low on the Likert scale, because they both have differences in principle, regardless of whether or not they are facially similar or different.
We now describe our students’ ratings to these four problems. We averaged the similarity ratings of each student for each problem pair type for each interview. Figure 5.27 below shows the rating for all four pair types for the first as well as second interview. The key on the top right of the figure is an abbreviated version of Table 5.2. It shows principle/facial similarities/differences in each type. (P=Principle, F=Facial, S=Similarity, D=Difference). The error bars are the standard deviation over all students and all problem pairs of a given type.

Figure 5.27 Students' similarity ratings of problem pairs of type A, B, C, and D for interview 1 and interview 2.

**Interview 1**

In interview 1 we find statistically significant differences between the similarity ratings of pairs A and B (p-value 0.000), B and C (p-value 0.003), and C and D (p-value 0.008). The fact that students have rated pairs B and D as significantly lower than pairs A and C is consistent with the notion that students appeared to be focusing on facial similarities/differences rather than similarities/differences in principle. For instance, they rated pair B significantly lower than pair A even though the problems in pair B were only facially different. Similarly, they rated pair C significantly higher than pair D even though the problems in pair C had differences in underlying principle.

Through discussion of the similarity ratings with students during this task, it became apparent that students recognized that problems were related by conservation of
energy, but they believed the differences in facial features have a direct effect on the types of energies involved, and these differences were enough to make the solution significantly more different.

Student EJ 1 “I guess that both the stone and the piano have potential energy like when they’re starting but that doesn't matter really, it's a totally different technique used to solve each problem. There a spring energy now.”

Student MD 1 “Ok these ones, this one has the spring constant that you use and you don't have this information here, it's a lot difference.”

It was also apparent through the conversation that Pair C problems were perceived by the students as different in terms of the method necessary to solve the problems, but are not ‘significantly’ different.

Student EJ 1 “Except this one you're gonna be using a tiny different equation in the path [solving procedure] than this one and that [part of the solution] was the same.”

Student MR-1 “And this one [problem] is kinda more complicated [than the other problem in the pair] because you're putting the friction coefficient and the work he is doing to counter that and both work to counter that to slow it down at the very bottom… so i mean they are pretty similar, just this one has to factor in the work that it have actually in ... i mean basically they are the same problem, as far as being identical it's the same problem.”

Further, it was common for students to mention the equation as a means for describing similarities between problems:

Student AM 1 “here your given mass and asked for spring constant, and here your given spring constant and asked for mass, and it has the same compression and both, and so youd use the exact same equation to solve for both, you would just have a different variable that you're solving for.”
Interview 2

In interview 2, we found that the differences between A and B, B and C were no longer statistically significant. The only statistically significant difference was between C and D (p-value 0.014). The fact that students were rating pairs A and B at about the same level of similarity is consistent with the notion that students have now begun to recognize that the problems in pair B have principle similarities that overpower their facial differences to the extent that they rated pair B almost the same way as they rated pair A in which the problems have both facial and principle similarities. In other words, it appears from these data that students were focused on the similarities in principle although there may be facial differences between the problems in pair B. The ratings for pairs C and D in interview 2 were statistically identical to their ratings for these pairs in interview 1. Particularly, we would be interested in seeing the rating for pair C to be significantly less than before, and as low as the rating for pair D. Such data would have been consistent with the notion that students are able to overlook the facial similarities in pair C and recognize the difference in principle. Our data do not appear to show this pattern. Rather, it appears from our data that when shown a problem pair that is facially similar, students do not probe further to reflect on whether or not these problems are similar or different in principle.

From interview discussion, it appeared that students continued to reflect on the facial differences and structural similarities, but the emphasis on the structural similarities with respect to devising a solution is different.

Student AM-2 “They are different… like the spring versus the pendulum. They're different on what you've asked for, but again because it's Simple Harmonic Motion it's the same kinda idea.”

Student DM-2 “This one I ranked pretty high because they are both …it's most likely gonna be pretty much the same but they are a different kind of harmonic motion….the equations are not exactly alike because you do have a spring stiffness [in the second problem], but they’re similar.”

Similarly to Interview 1, students continued to rank similarity in Pair C high, and it was interesting to hear their reflection on their own ranking. It would seem that for some
students the problems’ differences in complexity were not enough to call them ‘conceptually different’ and so their similarity did not drop far from Pair A. Other students were focused on the similarity in situation.

Student EJ-2 “So they’re asking for two different values at the end but for the most part you have to use most similarish equations, and maybe more for this one than the other one…its the same situation.”

Student DM-2 “I think all of them you can solve pretty similarly but this one you have to add stuff [points to equations on equation sheet] to it.”

Student SS-2 “So that would be two extra steps and so, they were fairly close, because they were the same method up to those two points.”

Summary

In summary, it appears that after completing all eight weeks of the focus group learning interview, the students in our cohort group were able to discern the similarities in principle between two problems in a pair that had facial differences and regard such similarities as important to solving the problem. But, given a pair with two problems that had facial similarities, they were unable to discern the differences in principle. This behavior appears to be consistent with the activities that they engaged in during the focus group learning interviews. Each week, problems were all focused on a single principle. We did not have problems in a given week that had any differences in principle. The only differences between the problems were facial differences. Therefore, the students appear to have developed the ability to use the facial differences as a cue to look deeper at a problem pair and decide whether there are any differences in deep structure, i.e. principle differences. This is consistent with pairs of type B being rated highly similar in interview 2. If the problems had facial similarities, however the students appear not to look deeper to ascertain whether or not the problems have similarities in principle. The students appear to decide, based on the facial similarities, that the problems are highly similar, without attending to the underlying similarities/differences in principle. This is why pairs of type C were rated highly similar in interview 2.

An important caveat in interpreting these data should not be overlooked. In our attempt to ensure that the problems presented to students in the problem pairs were on
topics that the students had covered most recently, we used the problem pairs in Figures 5.23 through 5.26 for interview 1. All of these problems were on energy conservation or the work-energy principle. Similarly, in interview 2 we used problems on the topic of simple harmonic motion that students had covered most recently in class. The differences observed in student ratings in interviews 1 and 2 could be attributed not just to the change in students’ ability to discern the similarities and differences due to participation in the focus group learning interviews, but they could also be attributed to the differences in the specific problems used in each interview, the topic on which they were based, or on the fact that students were also enrolled in the class during the semester which also could have improved their abilities on these problem similarity rating tasks.

To isolate the effect of the focus group learning interviews on students’ performance on the similarity rating tasks, we would need to complete interviews with students who were enrolled in the class but who did not participate in our focus group learning interviews. We would also need to have used the same set of problems for both interviews to eliminate the possibility that the effects observed are due to the specific problems being used in the interview and not the change in the students’ abilities between interview 1 and interview 2.

5.4.5 Usability Task – Methodology

For the final task, students were presented with a challenging problem and asked to predict which of the problems seen in the usability task would be most and least useful as a solved example to enable them to solve the challenging problem. The challenging problem for interview 1 is shown below in Figure 5.28. Figure 5.29 displays the challenging problem for interview 2.
Figure 5.28 Challenging problem for individual interview 1.

![Figure 5.28](image)

A ball of mass 2kg is held against a spring compressed by 1 m with a spring constant of \( k = 3 \times 10^5 \text{N/m} \), and sits at the bottom of a frictionless ramp 50 m high inclined at an angle of 60°. After leaving the ramp, the ball travels along the track of a frictionless rollercoaster. After leaving the rollercoaster the ball sinks in the sand to a depth of 1m. What is the average force on the ball by the sand?

Figure 5.29 Challenging problem given for individual interview 2.

![Figure 5.29](image)

The length of a simple pendulum is 0.80 m, the pendulum bob has a mass of 0.30 kg, and it is released at an angle of 12° to the vertical. At the lowest point of the swing, the bob strikes a 0.60 kg block attached to a fixed horizontal spring whose spring stiffness constant is \( 7.70 \times 10^5 \text{N/m} \). The block is set into vibration.

(a) What is the speed of the bob before it strikes the block?

(b) What is the resulting amplitude of the block after being struck by the bob?

(c) At what angle would we have to start the simple pendulum if our block is set into vibration with an amplitude of 0.20 m?

The interviewer would spread out the problems on the table, and students were instructed that they may move or pick up problems as they wished. Once a student selected the two problems they thought would be most useful and least useful to see fully solved out for assisting with the challenging problem, they were asked to explain why they chose the problem. Similar to all other tasks, students were broken up into two
separate groups, and each group was given a different problem set. The challenging problem remained the same for both sets.

5.4.6 Usability Task – Results

Interview 1

For set 1 in interview 1, three students selected problem 12 (see Appendix C-3 for problems given in the usability task for interview 1. Table 5.8 below includes problems selected as most useful in set 1, interview 1. Problem numbers are given in the lower right hand corner) as most useful, one student selected problem 11 as most useful, and one student did not complete the usability task for interview 1.
Table 5.8 Usability problems selected by students for set 1, interview 1 as most and least useful.

<table>
<thead>
<tr>
<th>Problem No.</th>
<th>Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td><img src="image" alt="Diagram of a piano being pushed" /> A 330 kg piano starts from rest and slides 3.6 m down to the bottom of a 28° incline while a man is pushing back on it parallel to the incline with a force of 400 N. Assuming there is no friction between the incline and the piano, what is the velocity of the piano at the very bottom of the incline?</td>
</tr>
<tr>
<td>5</td>
<td><img src="image" alt="Diagram of a piano being pushed" /> A 330 kg piano starts from rest and slides down to the bottom of a 28° incline while a man is pushing back on it parallel to the incline with a force of 400 N. Assuming there is no friction between the incline and the piano, and the piano hits the bottom of the incline with a velocity of 2 m/s, from what height did the piano start descending?</td>
</tr>
<tr>
<td>6</td>
<td><img src="image" alt="Diagram of a piano being pushed" /> A 330 kg piano starts from rest and slides 3.6 m down to the bottom of a 28° incline while a man is pushing back on it parallel to the incline with a force of 400 N. The man does 1440 J of work on the piano. The effective coefficient of kinetic friction between the incline and the piano is 0.40. What is the velocity of the piano at the very bottom of the incline?</td>
</tr>
</tbody>
</table>
Students selecting problem 12 determined that the problem did a good job at explaining spring energy. They believed that the extra steps required to complete the problem involving gravitational potential energy would also be useful to see. One student also commented that they would not ‘lose’ anything by selecting this problem over a different spring problem. When asked to clarify, he stated that he would be able to learn the same things about spring energy from all three problems involving springs, but this one also included an initial gravitational potential energy component, so it gave more information.

Student EJ-1 “most useful would be one of these spring ones, I just have to figure out which one of these I like best. This one actually, umm...because this I think its going to end up with potential energy and kinetic energy and spring energy and that would be nice to see solved out.”
The student who selected problem 11 as the most useful problem said that the problem would do a sufficient job at explaining how the spring’s energy could be described. When asked how they determined that problem 11 was better than any other spring problem, the students stated that they could have chosen other spring problems as well.

For the same set, two students chose problem 6 as least useful, one student selected problem 4 as least useful, and the one student that selected problem 11 as most useful, selected problem 5 as least useful. The students that chose problems 4, 5, and 6 as least useful all stated that the problem they needed to solve for required potential and kinetic energy, but they felt like that part of the calculation was ‘easy.’ They felt that their difficulties with obtaining a solution would lie with the spring, and so they all chose piano problems as being least useful. The students that chose problem 6 more specifically stated that the challenging problem did not have any friction, and so it would be unnecessary for them to see a problem that had friction and work, such as problem six. The student that chose problem 5 also stated that he disliked problems that involved work, but would prefer to see it solved so that he could understand it better. He felt like it would not hurt to see a problem like that. It was this reasoning that led him away from choosing problem 6 as least useful.

Student DM-1 “I already know how to do it, I understand potential and kinetic energy, so this would not be useful to me.”

Student MM-1 “I have trouble with the work-force and adding potential energy..i don’t know why but I don’t like that. I can’t say this one [problem 6], because it has to do with work, and if I saw this worked out, maybe it would help.”

For set 2 in interview 1, four students selected problem 9 as most useful, and one student selected problem 2 as most useful. See Table 5.9 for all problems selected by student for set 2, interview 1 as most and least useful.
Table 5.9 Usability problems selected by students for set 2, interview 1 as most and least useful.

<table>
<thead>
<tr>
<th>Problem No.</th>
<th>Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><img src="image1.png" alt="Diagram" /> An 800 kg roller coaster shown in the figure above is dragged up to point 1 where it is released from rest. Assuming the track is frictionless; calculate the speed at point 3.</td>
</tr>
<tr>
<td>2</td>
<td><img src="image2.png" alt="Diagram" /> A roller coaster shown in the figure above will be moving with a velocity of 22 m/s at the exact moment it hits point 2. Assuming the track is frictionless; calculate the speed at point 4.</td>
</tr>
<tr>
<td>8</td>
<td><img src="image3.png" alt="Diagram" /> A 0.10 kg bullet is loaded into a gun compressing a spring a distance of 0.20 m. When the trigger is pulled, the spring is released, and the bullet leaves the spring at the spring’s relaxed length. The bullet travels a distance of 0.60 m before exiting the barrel of the gun at a speed of 200m/s. The coefficient of kinetic friction between the bullet and the barrel is 0.10. What is the spring constant?</td>
</tr>
</tbody>
</table>
A 0.10 kg bullet is loaded into a gun tilted upward at a 30° angle from the horizontal, compressing a spring (spring constant is 6400 N/m) a distance of 0.20 m. When the trigger is pulled, the spring is released, and the bullet leaves the spring at the spring’s relaxed length at a speed of 50.5 m/s. The bullet travels a distance of 0.60 m before exiting the barrel of the gun. What is the speed of the bullet as it leaves the gun?

All four students that selected problem 9 determined that the problem was the best fit because it included a spring and an incline. The one student that selected problem 2 as most useful determined that the challenging problem given and problem 2 both started their calculation at the bottom point of a coaster and both ended with some velocity.

Student CF-1 “It has an incline, it has a lot of the same variables..uhh…and it has a distance also..yeah, it has all the same information except its missing the ball sinking. I can use this to solve most of the more complicated…one. Did I Pass?”

Student SS-1 “I think that ..i mean, they all start at the same points, but you don’t need to worry about the potential due to gravity because it starts down here and not at the top. They both have a velocity at the end..well until this one hits the sand...it will have a velocity.”

For the same set, three students selected problem 1 as least useful, and two students chose problem 8 as least useful, one of those two students being the one that selected problem 2 as most useful. Two of the three students that chose problem 1 as least useful stated that the problem was too easy and that they would be capable of solving that problem without any help. The third student cited problem 1 as least useful because it had the rollercoaster dragged up to the top where the challenging problem has a spring force. Two students that chose problem 8 both noted the kinetic friction given in the problem as
being an extra complicating factor. They decided it was best to make that problem least useful because they did not want to get confused by irrelevant information.

Student SS-1 “I don’t like when they have more information in them. Like this one has kinetic friction which is not necessary for helping me out. I think it would be extra confusing.”

It is interesting that there are only four distinct problems in each case that are classified as most and least useful. Given the small sample size it is not possible to determine that this result is statistically significant in any way, but it is important to note that problems 3, 5, 7, and 10 were never tagged most or least useful by any students during interview 1. More interestingly, there existed three problems in each usability set that used the same scenario. For interview 1, there was one scenario that always involved a spring, and another scenario that always involved an incline. When students selections of most and least useful problems were tabulated with regards to these surface features, it became apparent that students selecting spring problems as most useful, were likely to select an incline problem as least useful. One can see this trend below.

Table 5.10 Most and least useful problem selections categorized by surface feature in individual interview 1.

<table>
<thead>
<tr>
<th>Interview 1 (ENERGY)</th>
<th>SPRING 1</th>
<th>SPRING 2</th>
<th>SPRING 3</th>
<th>INCLINE 1</th>
<th>INCLINE 2</th>
<th>INCLINE 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOST Set 1 (# of people)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEAST Set 1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>MOST Set 2</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LEAST Set 2</td>
<td>2</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Interview 2**

For set 1 in interview 2, two students were unable to complete task 3 in the time given. One of these students was also unable to complete this task during interview 1. The three remaining students selected problem 21 (see Appendix C-4 for problems given in the usability task for interview 2. Problem numbers are given in the lower right hand
corner) as most useful. Table 5.11 displays problems selected by students as most and least useful for set 1, interview 2.

**Table 5.11 Usability problems selected by students for set 1, interview 2 as most and least useful.**

<table>
<thead>
<tr>
<th>Problem No.</th>
<th>Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>A common device for entertaining a toddler is a jump seat that hangs from the ceiling. Assume that jump seats’ elastic cords act like a large spring of stiffness constant $60 \text{ N/m}$. Once the 15 kg toddler is placed in the seat, a caregiver pulls down the jump seat an extra distance of 0.50 m from equilibrium and releases. Find the period of motion and the maximum velocity of the toddler.</td>
</tr>
<tr>
<td>9</td>
<td>A common device for entertaining a toddler is a jump seat that hangs from the ceiling. Assume that jump seats’ elastic cords act like a large spring. Once the 15 kg toddler is placed in the seat, a caregiver pulls down the jump seat and releases, allowing the toddler to oscillate up and down taking 3 seconds to complete one oscillation. The toddler never moves faster than 0.20 m/s. Find the maximum distance the toddler moves away from equilibrium and spring stiffness constant of the elastic chords.</td>
</tr>
</tbody>
</table>
Students selecting problem 21 all stated that the problem was selected because it asks for a cable angle with respect to the vertical and the challenging problem also requires a pendulum angle to be found. When asked to clarify whether the angle calculation would be the same procedure or a similar procedure, two students stated that the angle calculated for the challenging problem would the same with different numbers because they would both be maximum angles, while the third student stated that the calculation would be different because the problem selected is calculating an angle at some different time. The interviewer asked the third student to clarify why they had not chosen problem 23 instead of problem 21 as problem 23 actually required an angle calculated at some point along its trajectory that was not a maximum. The student stated that it was because problem 23 was more mathematically complicated than problem 21 because it calculated a velocity first.
Student CF-2 “This is asking for a specific angle at a point in time, and this one is asking for an angle at a max, and they’re not the same but if I see this, it will help me do this….problem 23 is more complicated, I think I could just use this one [problem 21] because I could find the velocity on my own.”

For the same set, two students chose problem 9 (not the same problem 9 as used for interview 1) as least useful, and one student selected problem 7 as least useful. All three students selected problems that involved the baby in the bouncy chair because they were spring systems. All three students stated in some way that the problems involving springs were less useful because there was only one question asked about the block and spring system in the challenging problem, but two questions regarding the pendulum.

The students that chose problem 9 as least useful determined that the problem was the least mathematically complicated problem of the three spring problems. Both students stated that the amplitude and spring stiffness constant were much easier to calculate than the velocity and displacement after some period of time. When both students were asked why they did not choose problem 7 over problem 9, one student stated that they could have chosen problem 7 for the same reasons, but chose problem 9 instead. The other student stated that problem 7 could be a little more helpful because a maximum velocity was calculated, and in simple harmonic motion, this should be a comparable calculation to a pendulum. When asked to describe how they were similar, the student then stated that he decided they were not similar, so he was not sure why problem 7 was not less useful.

The student that chose problem 7 as least useful stated that the problem was less helpful than the rest of the spring problems because it was not calculating the same quantity that was asked for in the challenging problem.

Student AM-2 “This is least useful because it has a different system, this one is a pendulum and a spring [challenging problem], and this one is just a spring [problem 7], and there’s stuff that is asked for in this one which might be helpful, but there are other problems that ask for the same thing that this [challenging problem] one asks for, so I would make this [problem 7] least useful.”
For set 2 in interview 2, two students selected problem 17 as most useful, while problem 5, problem 13, and problem 15 were selected by the remaining three students. See Table 5.12 for problems selected by students as most and least useful for set 2 of interview 2.

**Table 5.12 Usability problems selected by students for set 2, interview 2 as most and least useful.**

<table>
<thead>
<tr>
<th>Problem No.</th>
<th>Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td><img src="image1.png" alt="Problem 5 Image" /> A 0.40 kg mass is sitting on a 30° inclined frictionless surface and is attached to the end of a spring with a spring stiffness constant of 300 N/m. The other end of the spring is bolted to the top of the 3.0 m high incline. The mass is stretched 0.30 m from equilibrium and released. Find the position and velocity of the mass at t=5 seconds after it was released.</td>
</tr>
<tr>
<td>13</td>
<td><img src="image2.png" alt="Problem 13 Image" /> A pendulum consists of a 3.0 kg uniform disk fixed at the end of a 1.0 m long rod of negligible mass hanging from a frictionless pivot. The pendulum is tilted from its vertical position and released. The center of the disk is 0.30 m above its lowest swinging point prior to release. Find the time period of the pendulum and the maximum velocity of the disk.</td>
</tr>
</tbody>
</table>
Students who selected problem 17 stated that the problem was useful because it was solving for an angle in a pendulum. They also determined that the problem would be solving for an angle after some time $t$ had passed, and that this scenario was different from the one given in the challenging problem. When asked why they did not choose problem 15 instead, one student stated that problem 15 solved for extra unnecessary information regarding the length of the rod, while the other student stated that problem 15 was simpler and she would prefer to see a more complicated problem because it might help more with other problems.

Student SS-2 “I don’t know…it looks more similar….I would prefer to see the problem that is harder only because I can better understand the steps. I think they would be more, and I could use the problem to help with other problems later.”
The student who chose problem 15 as most useful stated that the problem asked for the same thing that was asked in the challenging problem. When asked why the angle was chosen as most important for them to see, the student stated that they thought they would be capable of solving for the ball’s velocity on their own, and that angles were always confusing. When asked why they chose problem 15 over problem 17 the student responded that problem 17 seemed to be a different calculation than what was being asked in the challenging problem.

Student AR-2 “It might not be solving for the same thing. Hmm… Actually, this one [problem 17] is solving at a certain time, so it’s different.”

The student who chose problem 13 as most useful stated that they were having some difficulty with this material and that problem 13 would be a great way to figure out how to do part (a) of the challenging problem. This student decided that they would worry about only one part at a time, and that problem 13 would be the best at getting them started.

Student MM-2 “I don’t really grasp this stuff like I did before this last exam. I think an easier problem would be better for me to learn from first. This earlier guy would be good to start with [problem 13].”

The student that chose problem 5 as most useful determined that their trouble was mostly with springs and that they would need to see one of the spring problems worked out. They also stated that problems that asked for velocities and positions at particular points in time were usually more difficult, and so they wanted to choose the most difficult spring problem to see worked out in the hope that it would make other spring problems easier.

Student MS-2 “This is more difficult than that one. I want to see the more difficult one worked out.”

Interviewer “What was it about the problem that gave you insight into the difficulty level?”

Student MS-2 “It asks for the velocity at some time…and the position…I think that makes it more difficult, mathematics-wise.”

For the same set, three students selected problem 5 as least useful, where all three students selected different problems as most useful. All three students who selected
problem 5 determined that the calculation of position and velocity of the mass at some particular time was more information than was needed to help with the challenging problem.

Student DM-2 “And this one is looking at the block system at a particular time, which is not like our problem because [in our problem] after the pendulum hits it is like the starting point of the spring problem. The other block problems would be better because they are looking at the block system from the maximum points.”

One student chose problem 1 as least useful and cited that it was not looking for a position such as the challenging problem.

Student MM2 “Uh..this one’s closer to me and I didn’t want to reach farther. Haha. Ok so these systems are the same, but this one is actually solving for amplitude like our problem, so its just a little more similar, and this ones solving for a position, so..more similar..yeah..this one’s [problem 1] least similar.

The remaining student chose problem 17 as least useful. The student who chose problem 17 as least useful was also the same student that chose problem 5 as most useful. The student stated that problem 17 was least useful because it was finding an angle at some “random time” and that this was not the same as finding the maximum angle. The student also decided that the pendulum part of the problem was easier than the spring part and so it made sense to pick a pendulum problem for the least useful.

Problems 3, 11, 19, and 23 were never tagged most or least useful by any students during interview 2. It is interesting to see that during interview 1 and interview 2, students continue to focus on the particular quantities that are asked to be solved for in a given problem. It is also often the case that the type of motion, like spring motion or pendulum motion, can have certain comfort levels associated with them. If a student has decided that a particular type of problem is difficult, that perception of difficulty will have an impact on why a particular problem is chosen. Thus, problems chosen as most useful in this final task are not necessarily picked because they are most similar to the challenging problem, but because they address students’ perceived difficulty.
It is also interesting to note that students seem more willing to learn from a more challenging problem solution by interview 2. During interview 1, students often tried to rid themselves of problems perceived to have higher levels of difficulty. During interview 2, there are a couple students that actually make selections of most useful problems that they perceive as more difficult because they hold more information and may be useful for assisting with a larger sampling of difficulties.

When problem selections were categorized by the surface feature found in the given scenarios, it became apparent, such as in interview 1, that student ratings were surface feature dependent. In other words, students selecting spring problems for most useful, were likely to select an incline problem as least useful. Each box in Table 5.13 below lists the number of people that chose a given interview problem as most or least useful. The problems are re-labeled by their most prominent surface feature to emphasize the student selection pattern. Also note that an asterisk is used to identify the one and only student that chose an incline problem as most useful and a spring problem as least useful.

**Table 5.13 Most and least useful problem selections categorized by surface feature in individual interview 2.**

<table>
<thead>
<tr>
<th>Interview 1</th>
<th>SPRING 1</th>
<th>SPRING 2</th>
<th>SPRING 3</th>
<th>INCLINE 1</th>
<th>INCLINE 2</th>
<th>INCLINE 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOST Set 1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(# of people)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEAST Set 1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>MOST Set 2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td>1*</td>
</tr>
<tr>
<td>LEAST Set 2</td>
<td></td>
<td>1*</td>
<td>1</td>
<td></td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

### 5.5 Non-Traditional Problems on Exams

Data were also collected from five multiple choice examinations taken during the semester. Individual scores for each examination question were obtained by the primary course instructor. It is important to note the timing of the exams relative to the focus group learning interviews. Figure 5.30 below shows when the exams were scheduled along the timeline. As seen below exams 1 and 2 were before the protocol of our group
learning interviews was finalized. Exam 3 was after the group learning protocol was mostly finalized, i.e. we had included the solved example (Problem C) into the protocol, and were providing structure in the protocol that would facilitate students to compare and contrast various problems and also reflect on the usefulness of Problem C while solving Problems A and B. Exams 4 and 5 were given after the protocol had been finalized. The difference between the finalized and ‘mostly’ finalized protocol is that rather than go over Problem C as we did in week 4 of the group learning interview, in group learning interviews 5 through 8 we presented Problem C to the students and asked them to go through it themselves and ask us any questions that they had about Problem C.

**Figure 5.30 Timeline showing when exams were administered during the semester relative to the FOGLI’s.**

![Timeline](image)

The last three problems on each examination were adaptations of text editing (Low and Over, 1990), physics jeopardy (Van Heuvelen and Maloney, 1999), and problem posing tasks (Mestre, 2002). While these tasks in the original form are open-ended, the problems included on the exams were in multiple-choice format for two reasons: first they conformed to the format of the rest of the test questions and second they could be graded efficiently for large numbers of students. We acknowledge that the open-ended tasks can provide richer information about the students’ conceptual knowledge, but we were content with the information of students’ conceptual schema provided by the multiple choice adaptations.

Text editing tasks, as described previously, involve presenting a student with a problem statement and then asking the student to identify the missing, irrelevant, and
required information in the problem statement without first solving the problem. Low and Over (Low and Over, 1990) point out that text-editing tasks can be a measure of schematic knowledge because they require an understanding of the deep structure of the problem. Because students are asked to complete the tasks without solving the problem, students need to know the interrelationships between various physical quantities, not in terms of equations, but at a conceptual level to be able to successfully complete the task. Figure 5.31 below shows an example of text editing used on one of the class exams.

**Figure 5.31 Example of a multiple choice adaptation of a text editing task.**

[Image of a multiple choice question with a diagram of a mass attached to a cord that passes over a pulley to another mass, with questions asking for the final velocity of the mass and options for not relevant quantities]

Physics Jeopardy tasks were first developed by Van Heuvelen and Maloney (Van Heuvelen and Maloney, 1999). As the name indicates, these tasks require the students to work backward. Students are given a fragment of a solution to a problem and asked to identify the physical scenario that corresponds to the solution. The developers point out that these tasks require an effort to represent a physical process in a variety of ways. Because of these features, students are unable to use naïve problem solving strategies while solving Jeopardy problems.

Figure 5.32 below shows an example of our adaptation of a Jeopardy problem that provides students with a few steps of a projectile motion. Students are asked to determine what trajectory shown corresponds to the problem. This task requires students to relate information given in the mathematical and symbolic representation to a visual or pictorial representation.
Problem posing tasks were used by Mestre and others (Mestre, 2002) in the context of physics problems. In the tasks presented by Mestre, students were given a scenario, typically in the form of a picture and were asked to construct a problem around the scenario that was based on certain physical principles. Mestre points out that problem posing tasks are aimed at probing students’ understanding of concepts as well as assessing whether they transfer their understanding to a new context. Clearly such a task was rather open-ended with multiple possible answers.

Our adaptation of this task is much more focused than Mestre’s original open-ended task. It presents students with the first part of a problem statement which clearly describes a physical scenario. Students are then asked to select from a list of choices, a question, which when added to the statement will create a solvable problem that requires the use of a set of given equations. Clearly, our adaptation differs significantly from the original problem posing task designed by Mestre. First, this task clearly does have a unique correct answer. Second, it requires specific conceptual knowledge, represented in the form of equations. An example of our adaptation of a problem posing task is shown in Figure 5.33 below.
Figure 5.33 Example of a multiple choice adaptation of a problem posing task.

You are given the starting statement of a problem below.

A 500 kg cargo shipment, attached to a parachute, drops vertically out of a helicopter hovering 100 m above a large spring (k = 220,000 N/m). The cargo comes to rest when the spring compression is 0.50 m.

Which question, when added to the statement above, will make a solvable problem that requires ALL of the following equations to solve?

\[ W = Fd \quad W = \Delta KE + \Delta PE \quad PE_{\text{grav}} = \frac{1}{2} kx^2 \quad PE_{\text{grav}} = mg \quad KE = \frac{1}{2} mv^2 \]

(a) What is the speed of the cargo just before striking the spring?
(b) How much time does it takes for the cargo to make contact with the spring?
(c) What is the work done by air resistance acting on the parachute as it drops?
(d) What is the average force of air resistance acting on the parachute as it drops?
(e) None of the above.

5.5.1 Results - Non-Traditional Problems on Exams

To assess students’ conceptual schema in problem solving we inserted three non-traditional problems on each of the five class exams during the semester. Each exam included a text editing, physics jeopardy, and problem posing task at the end. These problems were assigned for extra credit and presented in a multiple choice format similar to the rest of the exam. All results from the statistical analysis on the examination data may be found under Appendix E.

Students were given the opportunity to drop any one of five examinations and there were a large number of students that took advantage of this opportunity. Since students were able to throw out or not take one of their five examinations, any analyses which required the mean averages of all five examinations would result in the loss of 47 of our total 274 student sample. With the assistance of a statistician, it was determined that a problem-by-problem analysis for each individual examination would be appropriate, as the loss of participants was dependent only on the students that did not take a particular examination, not on whether they missed any one of five examinations. Our problem-by-problem analysis allowed us to determine whether there was a statistically significant difference between our treatment and baseline groups per problem type (traditional, problem posing, jeopardy and text editing) in each exam.

We investigated how the treatment and rest of the class varied in performance with respect to the overall averages from all five examinations on problem posing, physics Jeopardy, text editing, and traditional problems. This analysis did require
students to take all five examinations, and thus 47 class participants were removed from this analysis. We also investigated how course participants varied in performance between overall examination averages. This analysis also required all students to take all five examinations, and this analysis did not distinguish between the treatment group and the rest of the class.

The initial problem-by problem analysis is described directly below, followed by the exam-by-exam and exam-by-treatment interaction ANOVAs. It is also important to note that using a standard three-factor ANOVA in this case was not appropriate because it failed to model the correlation between the repeated measures. The data violate the ANOVA assumption of independence. Achieving independence would require separate participants taking separate exams, or even separate participants solving separate problems on each exam.

On each exam we compared the performance of our cohort group with the rest of the class on each non-traditional problem based on a logistics test using a binomial model. We also compared the performance of our cohort group with the rest of the class as on all of the traditional problems using an ANOVA single factor test $\alpha=0.1$ level of significance. The selection of our $\alpha$-level inevitably involved a compromise between significance and power, and as a result between the Type I error and the Type II error. The selection of the significance value is difficult, and often, $\alpha=0.5$ is selected simply out of tradition. In this sense, because our cohort group is a set of nine students (1 student participants scores were removed because they were given a different examination) and the rest of the course is over 200, any statistical analysis will yield moderate ‘significance’ levels. Thus at a standard $\alpha=.05$ level, we could narrow our window of opportunity for obtaining a false positive result, but that’s only assuming the nine students are a perfect representation of the variance within the whole of the population. Though they are a good representation of the variance, it would be difficult to argue that the variance of a 200+ class could be represented wholly by nine volunteers. In education research, it is not necessarily uncommon to see the $\alpha=0.1$ levels of significance used when the treatment samples are no larger than 10. In fact, with smaller sample sizes, there is often a delicate balance between an increase in power (or increase in level of significance) with respect to the sample size. The likelihood of correctly rejecting a null hypothesis is much lower for
smaller samples since the standard error is much higher given N<20, thus decreasing the accuracy of estimates of parameters (Snedecor and Cochran, 1989; McCall, 2001). Because of the arbitrary nature of how these alpha levels are selected, the results presented in our analyses will report direct probabilities or p-values along with whether they stand as significant or insignificant given our level of significance selection. Our goal for selecting the ANOVA was to determine whether there might exist positive differences, and this test, even at a lower significance level, will accomplish that goal. If we failed to see significance at $\alpha=0.1$, then we would have learned that the likelihood of the treatment doing any good would be very small. Even with a higher significance level, the ‘significance’ does not designate whether the difference is large or important, only that it has a certain probability of existing.

It is important to recall (Figure 5.26) that the first three exams were given before week 5 of the focus group learning interview during and after which the finalized protocol was used. Table 5.14 below shows the comparison of performance on traditional exam problems between our cohort group with the rest of the class on each exam, showing the mean and standard error on each exam. We find that there is no statistically significant difference (at the $\alpha=0.1$ level of significance) between our cohort and the rest of the class on their performance on traditional exam problems.

**Table 5.14 Comparison of scores of cohort and rest of class on traditional exam problems.**

<table>
<thead>
<tr>
<th>Exam #</th>
<th>Group Int. Cohort Mean ± S.E. (N)</th>
<th>Rest of the Class Mean ± S.E. (N)</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>75.3% ± 6.03% (N = 9)</td>
<td>70.0% ± 1.09% (N = 274)</td>
<td>0.3808</td>
</tr>
<tr>
<td>2</td>
<td>62.2% ± 6.11% (N = 9)</td>
<td>61.1% ± 1.08% (N = 274)</td>
<td>0.8559</td>
</tr>
<tr>
<td>3</td>
<td>69.7% ± 6.22% (N = 9)</td>
<td>65.0% ± 1.14% (N = 267)</td>
<td>0.4593</td>
</tr>
<tr>
<td>4</td>
<td>76.8% ± 4.98% (N = 9)</td>
<td>77.0% ± 0.93% (N = 258)</td>
<td>0.9795</td>
</tr>
<tr>
<td>5</td>
<td>79.4% ± 5.99% (N = 7)</td>
<td>77.6% ± 0.99% (N = 258)</td>
<td>0.7655</td>
</tr>
</tbody>
</table>

NONE are ≤ 0.10
Table 5.15 below shows the comparison of performance on text editing tasks between our cohort group with the rest of the class on each exam, showing the mean and standard error on each exam. We find that there is no statistically significant difference (at the $\alpha=0.1$ level of significance) between our cohort and the rest of the class on their performance on text editing problems.

Table 5.15 Comparison of scores of cohort and rest of class on text editing tasks.

<table>
<thead>
<tr>
<th>Exam #</th>
<th>Group Int. Cohort % Correct (N)</th>
<th>Rest of the Class % Correct (N)</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>44.5% (N = 9)</td>
<td>35.0% (N = 274)</td>
<td>0.5673</td>
</tr>
<tr>
<td>2</td>
<td>77.8% (N = 9)</td>
<td>74.1% (N = 274)</td>
<td>0.8003</td>
</tr>
<tr>
<td>3</td>
<td>55.6% (N = 9)</td>
<td>61.8% (N = 267)</td>
<td>0.7072</td>
</tr>
<tr>
<td>4</td>
<td>44.5% (N = 9)</td>
<td>44.6% (N = 258)</td>
<td>0.9339</td>
</tr>
<tr>
<td>5</td>
<td>42.9% (N = 7)</td>
<td>47.3% (N = 258)</td>
<td>0.3354</td>
</tr>
</tbody>
</table>

Table 5.16 below shows the comparison of performance on physics jeopardy tasks between our cohort group with the rest of the class on each exam, showing the mean and standard error on each exam. We find that there is no statistically significant difference (at the $\alpha=0.1$ level of significance) between our cohort and the rest of the class on their performance on physics jeopardy tasks, except on exam 5 when the students in our cohort group performed significantly better than students in the rest of the class (p value = 0.0635)
Table 5.16 Comparison of scores of cohort and rest of class on physics Jeopardy tasks.

<table>
<thead>
<tr>
<th>Exam #</th>
<th>Group Int. Cohort % Correct (N)</th>
<th>Rest of the Class % Correct (N)</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>55.6% (N = 9)</td>
<td>52.9% (N = 274)</td>
<td>0.8760</td>
</tr>
<tr>
<td>2</td>
<td>100% (N = 9)</td>
<td>92.3% (N = 274)</td>
<td>0.2348</td>
</tr>
<tr>
<td>3</td>
<td>55.6% (N = 9)</td>
<td>58.4% (N = 287)</td>
<td>0.8639</td>
</tr>
<tr>
<td>4</td>
<td>44.5% (N = 9)</td>
<td>33.7% (N = 258)</td>
<td>0.5127</td>
</tr>
<tr>
<td>5</td>
<td>100% (N = 7)</td>
<td>77.9% (N = 258)</td>
<td>0.0635</td>
</tr>
</tbody>
</table>

Only Exam 5 is ≤ 0.10

Table 5.17 below shows the comparison of performance on problem posing tasks between our cohort group with the rest of the class on each exam, showing the mean and standard error on each exam. We find that there is no statistically significant difference (at the 0.1 level of significance) between our cohort and the rest of the class on their performance on problem posing tasks except on exam 4 and exam 5 when the students in our cohort group performed better than students in the rest of the class (p value = 0.0012 on exam 4 and 0.0821 on exam 5 respectively).
Table 5.17 Comparison of scores of cohort and rest of class on problem posing tasks.

<table>
<thead>
<tr>
<th>Exam #</th>
<th>Group Int. Cohort % Correct (N)</th>
<th>Rest of the Class % Correct (N)</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>22.3% (N = 9)</td>
<td>34.7% (N = 274)</td>
<td>0.4226</td>
</tr>
<tr>
<td>2</td>
<td>22.3% (N = 9)</td>
<td>36.1% (N = 274)</td>
<td>0.3741</td>
</tr>
<tr>
<td>3</td>
<td>11.2% (N = 9)</td>
<td>21.7% (N = 267)</td>
<td>0.4117</td>
</tr>
<tr>
<td>4</td>
<td>88.9% (N = 9)</td>
<td>36.4% (N = 258)</td>
<td>0.0012</td>
</tr>
<tr>
<td>5</td>
<td>57.2% (N = 7)</td>
<td>25.6% (N = 258)</td>
<td>0.0821</td>
</tr>
</tbody>
</table>

Based on the data above one can see that students in our cohort group performed better than the rest of the class on two of the three non-traditional tasks (problem posing and jeopardy) on exam 4 and exam 5. The following aspects of these results are noteworthy.

First, there was no statistically significant difference between our cohort and the rest of the class on traditional problems on any of the exams. So, participating in the group learning interviews apparently did not improve the performance of our cohort group on traditional problems. These data are consistent with the notion that traditional problems are amenable to novice problem solving strategies and therefore are not effective assessment tools for gauging improvements in students’ conceptual schema in problem solving.

Second, the only statistically significant differences in the data above occurred in exam 4 and exam 5, which occurred after week 4 of the focus group learning interviews (See timeline displayed in Figure 5.30). These data are consistent with the fact that it was only after week 4 in the focus group learning interviews that we implemented the finalized protocol that explicitly required students to rank and describe the similarities and differences between problems A/B and A/C and B/C. Before week 4, students were not being provided with adequate procedural scaffolding to free up their cognitive
resources to engage in reflection about the similarities and differences between the problems.

We also compared the exam-by-exam and exam-by-treatment interactions using data from only those students who completed all of the exams. These analyses resulted in a loss of 47 participants per semester out of a total of 274 participants per semester and a loss of two of our nine students in the treatment group. As previously stated for the problem-by-problem analysis described above, both the normality and the constant-variance assumptions required for an ANOVA have not been violated.

In order to complete this analysis, each examination was divided into four types of problems: traditional examination problems, problem posing problems, text editing problems, and physics Jeopardy problems. Average scores, or mean scores, were calculated for each problem type for each examination. The exam-by-treatment interaction analysis determined whether there were differences between our baseline and treatment groups with respect to mean scores and differences in instruction between our baseline and treatment groups. The exam-by-exam interaction analysis determined whether there were differences in all class participants (baseline and treatment students) mean scores between examinations. Tables 5.18 through 5.20 below contain the p-values and description of whether differences are significance with respect to each problem type.

Table 5.18 below displays whether there is a statistically significant difference between the baseline and treatment groups with regards to mean scores on traditional and non-traditional problems. This is different from our original analysis as it uses the average score on all five examinations for a given problem type, not individual scores on individual examinations. At an alpha level of 0.10, we observe statistical differences between the baseline and treatment group scores on the physics Jeopardy task. All other problem types show no statistical differences between groups on their average scores. In order to achieve significance for this analysis, students in the treatment cohort would have to show a large improvement in the last three out of the five examinations such that they overcame the large standard error from their small population. Significant improvement on the first examination would not be expected because it occurred prior to the start of any treatment. Any significant improvement observed on examination two would be unusual given that the examination occurred after only two, 75 minute
treatments. Significance could also be achieved if students overall means were higher on all examinations in one group than the other. This is not likely with this sample of participants given that the significant differences occurred in only one of four problem types.

**Table 5.18 Significance between the baseline and treatment groups on traditional and non-traditional problem scores averaged over all five examinations.**

<table>
<thead>
<tr>
<th>Problem Type</th>
<th>p-value</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>0.4776</td>
<td>Not Significant</td>
</tr>
<tr>
<td>Jeopardy</td>
<td>0.0522</td>
<td>Significant</td>
</tr>
<tr>
<td>Text Editing</td>
<td>0.9552</td>
<td>Not Significant</td>
</tr>
<tr>
<td>Problem Posing</td>
<td>0.5925</td>
<td>Not Significant</td>
</tr>
</tbody>
</table>

Our exam-by-exam interaction analysis indicated significant differences (at an alpha level of \(\alpha=0.10\)) in performance on all problem types between the five examinations. This significance suggests that the average scores for different problem types vary significantly between examinations. Table 5.19 displays the p-values (significance of differences) in average problem scores on traditional and non-traditional problems between examinations.

**Table 5.19 Significance of difference on average traditional and non-traditional problem scores between examinations.**

<table>
<thead>
<tr>
<th>Problem Type</th>
<th>p-value</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>0.0001</td>
<td>Significant</td>
</tr>
<tr>
<td>Jeopardy</td>
<td>0.0001</td>
<td>Significant</td>
</tr>
<tr>
<td>Text Editing</td>
<td>0.0001</td>
<td>Significant</td>
</tr>
<tr>
<td>Problem Posing</td>
<td>0.0019</td>
<td>Significant</td>
</tr>
</tbody>
</table>

Using the split plots shown in Figures 5.34 through 5.37 below, we also observe that the differences in average scores between examinations were not always positive. That is, the average scores for a given problem type are statistically different between examinations, but the averages don’t always improve as the semester progresses. The exam-by-exam interaction analysis does not distinguish between the baseline and treatment groups, as it uses the average scores for the total class population, but we may use the univariate split plots below to uncover more information regarding performance differences between the treatment and rest of the class.
The univariate split plot analysis determined whether there existed a significant difference at $\alpha=0.10$ in instruction on traditional or non-traditional problem scores between the examinations. The primary difference between this analysis and our previous problem-by-problem ANOVA was that the univariate split plots investigate differences in problem type averages between examinations 1 through 5 while the problem-by-problem analysis investigated differences in instruction by comparing individual students’ problem type performances on individual examinations. Statistically significant differences between the baseline and treatment groups instruction on problem posing problems were observed through mean examination scores. There were no other problem types displaying significant differences. The one-factor, problem-by-problem ANOVA detailed previously also showed significant differences between the baseline and treatment groups for problem posing problems specifically on examinations 4 and 5, so our split plot analysis does not contradict our previous results. Table 5.20 displays the p-values (significance of difference) between instruction on traditional and non-traditional problem scores between examinations.

**Table 5.20 Significance in difference between the treatment and baseline groups on traditional and non-traditional problem scores between the five examinations.**

<table>
<thead>
<tr>
<th>Problem Type</th>
<th>p-value</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>0.9979</td>
<td>Not Significant</td>
</tr>
<tr>
<td>Jeopardy</td>
<td>0.5049</td>
<td>Not Significant</td>
</tr>
<tr>
<td>Text Editing</td>
<td>0.6311</td>
<td>Not Significant</td>
</tr>
<tr>
<td>Problem Posing</td>
<td>0.0022</td>
<td>Significant</td>
</tr>
</tbody>
</table>

We can observe treatment by examination interaction by plotting our mean scores across examinations between the baseline and treatment groups. Theoretically, the generated split plot lines should cross or cross eventually if extended when there is any interaction. Since our data are not error free, these plots are used as a graphical interpretation of our calculations shown in Table 5.20. Below, Figures 5.34 through 5.37 display the split plots for traditional and non-traditional problem scores. We can see that there was one case of significant crossover, or difference in trends between treatments across repeated measures. This case occurs in Figure 5.37 for problem posing and matches with the p-value shown in Table 5.20.
Figure 5.34 Split plot of the baseline and treatment groups average scores on traditional problems for each examination.
Figure 5.35 Split plot of the baseline and treatment groups average scores on Physics Jeopardy problems for each examination.
Figure 5.36 Split plot of the baseline and treatment groups average scores on text editing problems for each examination.
As before, an important caveat in interpreting these data should not be overlooked. The topical content of material covered on each of these exams was very different. The level of difficulty of the non-traditional problems and traditional problems on each exam was also very different. Therefore any differences between scores on traditional or non-traditional problems on exams could also be the result of these differences, rather than a result of the participation of our cohort group in the focus group learning interviews. There is also still the possibility that the ANOVA’s revealed false positives, though there would be no way to discern whether this was true without replicating our study with a much larger population of students. This charge would be too difficult to accomplish for our project and would require full integration with the course and instructor.

5.6 Summary

The goal of this study was to examine whether participation in appropriately designed learning activities could facilitate students’ development of conceptual schema
in problem solving through appropriate use of case-based reasoning. To achieve this goal we conducted a series of eight weekly focus group learning interviews with a cohort of 10 students in an algebra-based physics class over the course of a semester.

**How do students perceive the usefulness of an example problem given a new problem possessing the same physical principle but lack similar objects and/or orientations?**

In the last five weeks of the focus group learning interviews we presented students with a solved example before asking them to attempt an unsolved problem. We found that students look for procedural elements in the solved problem that map on to the unsolved problem, such as the sequence of steps followed and equations used. If students are aware of an easier procedure to solve a problem than used in the solved example, they tend to ignore the solved example completely. Even by the end of the study we found that students’ focus continues to be on procedural case reuse rather than schema abstraction.

**Given problems that are deep-structure similar and surface feature different, what similarities and differences do students offer as important to the problem solutions?**

Our results indicate that after the focus group learning interview protocol was finalized as described above, students were better able discern the principle similarities and differences between problems, only if the problems were facially different. If the problems were facially similar, it appears that students were not cued to look deeper and they appear not to have been able to discern differences in underlying principles. It may also be indicated from individual interview observations that student selection of examples may be heavily guided by their own comfort level with a given problem type with mixed associations between surface and deep-structure characteristics. Problems such as the ‘spring problem’ all carry the same surface characteristic of a spring, but are also designated the more difficult object to calculate things for with respect to simple harmonic motion. We might also infer from our individual interview observations that students focus on the quantity that is asked for in a problem more than any other quantity given in an example. This focus on the “asked for” quantity is consistent with novice “means-ends” analysis.
We also learned that it is important to choose the level of difficulty of the problems carefully to ensure that students focus not on the mechanics of solving the problem correctly, but rather reflect on the similarities and differences between various problems. The protocol of these group learning interviews evolved over the course of the semester such that these reflections were emphasized. In the finalized protocol we provided students with a solved example as well as a structure of eliciting the underlying principles and explicitly comparing and contrasting the problems. This procedural scaffolding appeared to have enabled students to relieve cognitive resources from the task of simply solving the problem correctly and focus instead on reflecting on the problems thereby facilitating the development of deeper conceptual schema.

**How does students’ emphasis on similarities and differences change given problem pairs with varying deep structure and surface feature similarities / differences?**

The students participating in our focus group learning interviews were able to discern the similarities in principle between two problems in a pair that had facial differences and regard such similarities as important to solving the problem. But, given a pair with two problems that had facial similarities, they were unable to discern the differences in principle between the two problems.

**How does students’ performance on problem solving differ with respect to the rest of the class?**

Our results also indicate that participation in the focus group learning interviews does not appear to improve students’ performance on traditional exam problems. But, on two of the three non-traditional exam tasks there is some statistical evidence of a difference between our cohort group and the rest of the class on specific examinations taken after the focus group learning interview protocol was finalized. Statistically significant differences in mean scores on problem types overall occurred only on one of the 4 types of problems, where the cohort group obtained a statistically higher mean average on problem posing problems.

### 5.7 Limitations and Future Work

In spite of these promising results it is important to note that in addition to the focus group interviews, students were continually studying different material in the class...
as the course progressed. Any improvement seen in our data could also be due to the
differences in specific exam problems, assigned homework problems, or individual
interview tasks, as much as it could be due to the effect of the learning interventions in
our focus group learning interviews.

These results by themselves are not enough to say that one should infuse their
class with case reuse strategies, but it does give us enough positive observations that we
would want to further continue this study. In the next chapter, we will describe a
replication phase, where focus group learning interviews are again conducted with
algebra based physics participants. It would have been preferable for statistical analysis
to obtain a larger sample of focus group participants, but it would be too difficult for the
two moderators/interviewers working on this project to take in any more than 12
volunteers in a given session.
CHAPTER 6 - Phase III – Replication of phase II: Focus Group Learning Interviews

6.1 Introduction

In our previous phase (Phase II), we conducted group learning interviews with a group of students enrolled in an algebra-based physics course. Students were asked to solve a physics problem using a detailed example problem. These example problems were chosen such that they shared physical principles, but not facial characteristics like similar objects or similar axial orientations. Students were then asked to individually and collaboratively reflect upon the structural and surface feature similarities and differences between physics problems. The intended focus of each interview was to reinforce expert-like problem comparison while slowly breaking down any pre-existing surface feature oriented associations common among novice students. Procedural scaffolding was added to minimize student emphasis on the process of solving a given problem.

The impact of our group learning interview intervention on students’ conceptual schema was assessed using non-traditional problems on examinations, similarity rating problems and usability ratings in individual interviews at the mid-point and toward the end of the semester, and problem posing tasks implemented at the end of each focus group learning interview. Trends resulting from the group learning interviews, the individual interviews, and the examinations were laid out in our previous chapter.

The results from Phase II were promising as they clearly showed some improvement in students’ ability to perform conceptual/principle-oriented tasks. These tasks, unlike traditional textbook style problems, required students to dispense with procedure-oriented problem solving approaches and focus on object interactions and measured quantities governed by the physical concepts and principles.

The study in Phase II was replicated in the following semester using another cohort of algebra based physics students. Given the small sample size, the large number of outside contributing variables (course instructor, homework, lecture, labs), and the initial trials associated with a first implementation, it was necessary to replicate our experiment to verify whether the results would hold.
Participants in this study (Phase III) included 12 students that were randomly selected from 46 volunteers. The participants met in a single group a total of six times during the semester for focus group learning interviews. Each of the six focus group learning interview sessions was 75 minutes long. Additionally, the participants met with the moderator individually for 50 minutes at the end of the semester. Figure 6.1 shows the research timeline beginning from left at the start of the semester to the right at the end of the semester.

Figure 6.1 Research timeline for the replication of the focus group learning interviews.

In this chapter, I will describe the methodology and results of the replication of the focus group learning interview study. The following research questions will be addressed:

- How do students perceive the usefulness of an example problem given a new problem possessing the same physical principle but lack similar objects and/or orientations?

- Given problems that are deep-structure similar and surface feature different, what similarities and differences do students offer as important to the problem solutions?
• How does students’ emphasis on similarities and differences change given problem pairs with varying deep structure and surface feature similarities / differences?
• How does students’ performance on problem solving differ with respect to the rest of the class?
• How do our results from the previous focus group learning interview study (Chapter 5) align with this study?

6.2 Screening Interviews

Students were asked to volunteer for a semester long study requiring 1 hour and 15 minutes of their time each week. Forty six student volunteers signed up, and a total of 18 volunteers were selected at random to participate in screening interviews. The screening interviews, each lasting about 20-30 minutes, were used to determine whether students were legitimately interested in participating in the interviews, were interested in working in groups with others in their class, and unlikely to drop out of the study. The screening interview protocol was identical to the one used the previous semester. This interview was conducted in the first three weeks of the semester.

6.2.1 Results from screening interviews

Of the 18 interviewees, only five had not taken physics in high school. The students’ majors varied from pre-veterinarian/animal science to physical therapy. During the screening interview, students were asked to describe their study habits and problem solving procedure. Fourteen of 18 students stated they read the book either before lecture or right after lecture. Twelve of 18 students described the book as ‘simple’ and an ‘easy read.’ Two of 18 students described the textbook as ‘too wordy.’

Student “I read the introduction section and the little boxes on the sides before I go to class. I figure it could help me out with understanding the work we cover in class.”

Interviewer “Do you read past the introduction before you go to class, or after?”
Student “Sometimes. It depends. [scratches ear and pauses to stare at the
textbook] The author is too wordy so I don’t like to read more than I have
to.”

Four of 18 students stated they did not use the textbook for reading. Three out of four of
these students suggested that the book was only good for getting the homework problems
while the fourth student believed the text could be used for examples that help with
solving homework problems.

“No way..no way..not not reading this. I can use it with homework to find
examples that look like homework problems.”

The problem solving procedures described by students were not only similar to one
another, but identical to the procedures provided by students in Phase II. The steps
outlined by all students could be generalized as follows: read the problem, pick out
information that was given and asked for in the problem, solve the problem using one or
more formulas which contained the quantities identified in the previous step.

Student “After I read it, you know, I find the particulars in an equation from the
section, and then plug the numbers in.”

Interviewer “Do you often know what equation or equations are necessary?”

Student “Sorry? oh, yeah. So far so good.”

When students were asked how they might approach resolving a more difficult
problem, again, students reported similar procedures as last semester. All but two of the
18 stated they would look to the book or lecture notes for examples for help. The
outlying two students would not address the question.

Student “I haven’t had trouble with the problems.”

Interviewer “That’s good. What might you do if you did run across a problem
that you could not find an equation for right away?”

Student “I don’t know. I think I would still try my way until it worked.”

Finally, when students were asked how they might work with others in groups, all
participants stated that they would like to work in groups. Seventeen of 18 students
stated that they already worked with others while solving the homework problems. The
remaining student stated he would like to work with groups, but has not found someone
that can meet with him on his tight schedule.
“I think working in groups or with other people would be great. I learn well that way because I can listen to others and other people can learn from me. I would work in a group now except I have practice and work and so I don’t have a lot of time to meet with other people.”

Based on the screening interviews, we determined that all 18 students invited for screening interviews would serve well for our study. The 18 students were divided into two groups, one female and one male. In order to match the classroom diversity we randomly selected 8 women and 4 men from the two piles. Unfortunately, only five participants from this selection, 2 women and 3 men, accepted the offer to participate in the full semester long study. The remaining students from our screening interviews were contacted the following week. One of our participants also invited his fraternity brothers, also enrolled in the same algebra-based physics class, to participate alongside him. By the second week of our interviews, we had 15 volunteers. Twelve of these 15 continued to participate in our focus group learning interviews for the remainder of the semester.

Table 6.1 Demographics for the 12 participants in this study.

<table>
<thead>
<tr>
<th>Code ID</th>
<th>Major</th>
<th>Previous Physics Classes</th>
<th>M/F</th>
<th>Class yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>Animal Science/Pre-Veterinarian</td>
<td>High School Physics</td>
<td>M</td>
<td>3 yrs</td>
</tr>
<tr>
<td>AP</td>
<td>Nutritional Sciences</td>
<td>High School Physics</td>
<td>F</td>
<td>4 yrs</td>
</tr>
<tr>
<td>AS</td>
<td>Construction Science</td>
<td>None</td>
<td>M</td>
<td>2 yrs</td>
</tr>
<tr>
<td>AT</td>
<td>Agronomy</td>
<td>High School Physics</td>
<td>F</td>
<td>3 yrs</td>
</tr>
<tr>
<td>BL</td>
<td>Life Science</td>
<td>High School Physics</td>
<td>M</td>
<td>6 yrs</td>
</tr>
<tr>
<td>BR</td>
<td>Physical Science</td>
<td>None</td>
<td>F</td>
<td>4 yrs</td>
</tr>
<tr>
<td>GC</td>
<td>Pre-Vet</td>
<td>None</td>
<td>M</td>
<td>6 yrs</td>
</tr>
<tr>
<td>JF</td>
<td>Microbiology</td>
<td>None</td>
<td>M</td>
<td>4 yrs</td>
</tr>
<tr>
<td>JK</td>
<td>Animal Science/Pre-Veterinarian</td>
<td>High School Physics</td>
<td>M</td>
<td>2 yrs</td>
</tr>
<tr>
<td>JS</td>
<td>Physical Therapy</td>
<td>High School Physics</td>
<td>F</td>
<td>2 yrs</td>
</tr>
<tr>
<td>KG</td>
<td>Agronomy</td>
<td>High School Physics</td>
<td>F</td>
<td>3 yrs</td>
</tr>
<tr>
<td>SB</td>
<td>Kinesiology</td>
<td>Health Physics</td>
<td>M</td>
<td>2 yrs</td>
</tr>
</tbody>
</table>
6.3 Focus Group Learning Interviews

6.3.1 Methodology for Focus Group Learning Interviews

There were a total of six focus group learning interviews spread across the semester, and one individual interview conducted with each participant at the end of the study. Our previous study required students to participate in eight focus group learning interviews and two individual interviews. However, our cohort was unable to meet for the focus group learning interviews at any time other than the scheduled course examination time, in weeks when there was no exam. The focus group learning interviews were conducted during this time on all weeks except for dates coinciding with examinations. The mid-semester individual interview was also removed from the study as there was not enough time in the semester to complete both interviews. When students were given an examination in class, they were unable to meet with us. Thus, we were not left with sufficient time to complete focus group interviews and two individual interviews. It was decided we would use the single individual interview at the end of our study to simply assess whether our students emphasized deep-structure similarities within problems by the end of the semester.

During each focus group learning interview session, a moderator would hand out an example problem for students to study. Students were asked to read through the example and ask questions if they had difficulties understanding any part of the problem solution. Once students felt comfortable with the solution, each student was handed a worksheet with one of two unsolved problem on the front. Participants were paired together such that one would be asked to work on ‘problem A’ while the other worked on ‘problem B.’ All of the problems shared deep structure similarities but had surface differences. Deep structure similarities for various interview problems were physical principles that were recently covered in classroom lecture and homework. These included but were not limited to the work-kinetic energy principle, rotational motion using moments of inertia and torque, and simple harmonic motion. Surface features, such as the interacting objects, were often made as different as possible given the topic covered for the interview. An example of the example problem and problem pairs used during the third focus group learning interview are shown in Figure 6.2 below.
Figure 6.2 Example problem C and problems A and B from week 4.

**Problem C (Example Problem)**

Joshua pushes a 1.85 kg box along a flat horizontal table applying an average force of 39.0 N. The box starts at rest and reaches a velocity of 12.0 m/s. Neglecting friction, how far did Joshua push the box?

![Diagrams showing forces and motion](image)

We may express the work done by Joshua on the box in terms of two quantities:

\[ W = F \cdot d_{\text{covered}} \]  

i. The force applied on the box by Joshua

ii. The distance covered while Joshua applied the force

This expression (1) may be rewritten in terms of the magnitudes of the force and distance travelled.

\[ W = F_{\text{box}} \cdot d \cos \theta \]  

where \( \theta \) is the angle between the directions of the displacement and the force.

(Full Solution to Problem C available in appendix D-2)

**Problem A**

A 0.10 kg arrow is fired from a bow. The bow is pulled back a distance of 0.8 m so that the arrow is released with a speed of 50 m/s as it leaves the bow. The arrow travels 25.0 m before hitting its target. What is the average force exerted on the arrow by the bowstring?

**Problem B**

A Yankees batter hits a 0.14 kg baseball sending it off into left field, 40 m away from the batter’s box. The baseball lands in a Royals fielder’s glove, exerting an average force of 300 N, moving the glove backward 0.25 m before coming to rest. What is the speed of the ball just before it is caught?

All three problems presented in the Figure 6.3 below are focused on the same physical principle (work-kinetic energy), but have many surface differences such as the kinds of objects (arrow vs. baseball), and the points along motion (arrow leaving the bow vs. the baseball hitting the fielders’ glove).

The worksheet also asked students to perform a series of steps to alleviate the cognitive load associated with the procedural aspect of solving a problem. Our focus was to not to investigate how students procedurally worked through a problem, but to assist students in the process of building better mental organization of the relationships between quantities, concepts, and principles. Thus, the worksheet provided procedural scaffolding.
to lessen student focus on problem solving procedure. Also, the worksheet was designed to elicit student reflection on the concepts and principles that are shared among the three problems.

**Figure 6.3 Example of step procedure given during the focus group learning interviews.**

<table>
<thead>
<tr>
<th>Information in Problem</th>
<th>Required OR Irrelevant</th>
<th>How do you know?</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.20 kg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1 m</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7 (inside the equation of motion)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2π(3) f (inside the equation of motion)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is there any missing information? If so what is it?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For step 1, students were asked to reflect on the principles involved in their unsolved problem prior to planning a solution. Step 2 required students to hypothesize as to what information given in a problem might be irrelevant, if any. Steps 3 through 5 required students to work through the procedure of solving the problem by drawing a diagram, writing out full equations necessary for the solution, and finally plugging in values to resolve the answer.
Steps 6-11 asked students to both individually and collaboratively come up with a list of similarities and differences between their assigned problem and the example problem, then between their problem and their partner’s problem. Students then ranked the similarities and differences according to how important they were with regards to obtaining a solution. Students were also asked to rate the usefulness of the example problem in helping students to solve their assigned unsolved problem. The Likert style rating scale went from 1 to 5, where 1 was least useful and 5 was most useful.
Finally, after students completed steps 1 through 11, they were asked to work with their partner to create their own problem which used “elements from problems A, B, and C.” Students then exchanged their problems with other pairs. Students would attempt to solve each others’ problems. If the problems were not solvable, the pair given the unsolved problem would have to determine why the problem was unsolvable, whether it was under-specified or over-specified.

6.3.2 Focus Group learning Interview – Week 1

Methodology

For the first week of the focus group learning interviews, students recently covered forces in two dimensions. Problems A, B, and C were created to incorporate two dimensional force problems. Problem C incorporated kinetic friction and motion in two dimensions. Problem A and B involved motion in one dimension, while problem A also included kinetic friction. Figure 6.6 below shows the problems created for the interview. Problem C included a full detailed solution. It may be viewed in full in Appendix D-2. Only five of our original 12 students selected to participate in the focus group learning.
interviews showed up for the first week. Though students did not have difficulty with the tasks presented, the initial introductions took much of the interview time. Students were unable to complete tasks past solving their problems individually.

Figure 6.6 Problems used during FOGLI 1.

**Problem C: SOLUTION**
A box $m_1 = 28.0 \text{ kg}$, lying on a drafting table and is connected to another box $m_2 = 14.0 \text{ kg}$ by a cord running over a frictionless pulley. The coefficient of static friction between the table and the block $m_1$ is $\mu_{s1} = 0.450$, and the coefficient of kinetic friction between the table and the block $m_1$ is $\mu_{k1} = 0.320$. At the instant shown, block $m_2$ is moving downward at a speed of $0.05 \text{ m/s}$ and is $1 \text{ meter}$ above ground level. What is the acceleration of each block?

**Step 1:** Draw the free body diagram for each block showing the coordinate $x$ and $y$ axes.

**Step 2:** On the free body diagram for each bin step 1 indicate the forces acting on the block.

**Step 3:** Using the angles given, find the $x$ and $y$ components.

**Problem A**
Two crates, of mass $m_1 = 75 \text{ kg}$ and $m_2 = 110 \text{ kg}$, are in contact and at rest on a horizontal surface. A $630 \text{ N}$ force is exerted on the $75 \text{ kg}$ crate. If the coefficient of kinetic friction is $0.15$ and the coefficient of static friction is $0.20$, calculate the acceleration of the system.

**Problem B**
The two masses, $m_1 = 2.2 \text{ kg}$ and $m_2 = 3.2 \text{ kg}$, are each initially $1.80 \text{ m}$ above the ground, and the frictionless pulley is $4.80 \text{ m}$ above the ground. Determine the acceleration of the lighter mass.

---

**Results for Week 1**

**Principles** – Four of five students determined that the given problems involved forces and Newton’s second law. Two of these students were given problem A while the other two
were given problem B. The fifth student, given problem A, determined that the problem involved forces and friction.

Student JK “This is like the problems we solved in our last homework assignment. They involve forces on objects and accelerating blocks.”

**Text Editing** – All five students were able to recognize the irrelevant information given in the problem. Students given problem A recognized that the coefficient of static frictions was not needed. Students given problem B recognized that the distance to the ground from the pulley and the block were unnecessary for solving for acceleration.

Student KG “1.80 m, acceleration is not relative to distance, and 4.8 m is irrelevant in this situation.”

Student SB “0.2, we only need this if it's not moving.”

**Similarities and Differences** – Only two of five students were able to rate the similarity between their problem and problem C. As stated earlier, the initial introductions took much of the interview time and most students were unable to complete this task. Both students rated how similar problem B was to problem C. Both students gave the problem a similarity rating of ‘3.0.’ During group discussion, students determined that the problems were similar because masses were accelerating and both problems could be solved using Newton’s second law. Students also determined that the problems were different because the masses were not the same and problem C included motion in two directions. One of the five students noted that they felt this was an unimportant difference because all it changed was the substitutions in the formula, not the whole formula.

Student SB “Just change the substitution, not the whole formula. F=ma still works.”

Students were also asked to rate the similarity between problem A and problem B, but it was too late in the interview to begin another discussion so the students were dismissed.

**Usefulness Rating** – Students were unable to complete the usefulness rating. The initial introductions took much of the interview time and most students were unable to complete this task.

**Problem Posing** – Students were also unable to begin the problem posing task.
6.3.3 Focus group learning interview – Week 2

Methodology

For week 2 of the focus group learning interviews, problems A, B, and C covered circular motion. Figure 6.7 below shows the problems created for the interview. Problem C included a complete solution that is not shown below. It may be viewed in full in Appendix D-2. Our entire cohort showed up for this interview, though introductions were in order for new students. All tasks except for the problem posing task were completed before the end of the interview.

Figure 6.7 Problems used during FOGLI 2.

Problem C
A puck of mass \( m = 1.50 \text{ kg} \) slides in a circle on a table while attached to a hanging cylinder of mass \( M = 2.50 \text{ kg} \) by a cord through a hole in the table. The coefficient of static friction is 0.20 and the coefficient of kinetic friction between the table and the puck is 0.10. The distance from the puck to the hole is 0.02 m and 0.04 m from the hole to the cylinder. What frequency is required to keep the cylinder at rest?

Step I: Draw the Free-Body Diagrams for the puck and cylinder.

Problem A
A block of mass \( m = 75.0 \text{ kg} \) is sitting on a rotating disk as shown in the figure. The coefficient of static friction between the blocks and the turntable is 0.75 and the coefficient of kinetic friction is 0.60. The mass is 0.10 m from the center of the turntable and the turntable has a radius of 0.12 m. Determine the maximum frequency the turntable may rotate such that the blocks do not move from their current position.
Results for Week 2

Principles—Six of the twelve students determined the given principles involved were centripetal acceleration. Two of these six also state that static friction is involved. These students were both assigned problem A which requires the use of static friction. One of the six also explained that the concept was ‘still Newtonian principles.’

The other six of twelve students also state that centripetal acceleration is necessary, but did so while explaining their problem objective.

   Student JK “the block is supposed to stay still while the disk moves, we want to know how fast it can spin without the block moving m, how fast is the frequency… we need the friction and circle speed.”

   Student JS “I am supposed to find the maximum period of revolution necessary to keep the person from falling. You will use the radius & the frictions (static & kinetic) to determine the speed & revolutions to keep the person up.”

Text Editing—All students were able to determine the irrelevant information given in the problem statement. Students given problem A stated that the radius of the table and coefficient of kinetic friction are unnecessary because the block is not moving and the distance to the block from the center is important. Students given problem B determined that the coefficient of kinetic friction is not important because “the person is not moving.”

Unlike the previous semester, students were not struggling with this task.

Similarities and Differences—Students overall rated all of the problems as slightly similar to slightly dissimilar for week 2. Students given problem A, on average, gave problem C a similarity rating of 2.85. Students given problem B, on average, gave
problem C a similarity rating of 2.3. When students were asked to give their reasoning, students who were asked to solve problem A stated that the three top similarities (in order of ranking) between the problems were the use of static friction, the use of frequency, and the use of centripetal acceleration. Students who were asked to solve problem B stated that the top two similarities between the problems (in order of ranking) were the use of static friction and centripetal force.

Student AP “Both established what they wanted to determine, frequency; both used static friction, and the equation with centripetal acceleration used were the same.”

Students who were given problem A stated the problems were dissimilar because a tension force is not used. Students given problem B stated that the problems were dissimilar because “you are solving for different quantities.” One student given problem B stated that he was unsure of what the differences were between the problems, but initially rated his problem as being 2.5 out of 5 for similarity. If students communicated some similarities and differences, they were often left unquestioned. However, these students refused to answer questions regarding the dissimilarity between problem B and problem C.

Student JS “IDK.”

Interviewer “You rated the problems 2.5 out of 5 in similarity. Can you talk a little bit about the differences here?”

Student JS “uh no. haha. I don’t remember.”

Interviewer [stands ground for a moment]

Student JS “oh, they just aren’t identical, ok?”

Problem A and B were also rated in similarity by all student participants. The average rating of similarity between these problems was 2.33. The ratings varied from 2 to 3 between all the students and the variance was independent of the problem assigned to the student. Students’ top two similarities between problems were “the same basic formula applied” and “could use problem C.” Students’ top two dissimilarities between problems were different directional orientations and the fact that one of the problems dealt with maximum rotation to keep an object from falling while the other was to keep it from slipping.
Table 6.2 Similarity and Usability Ratings for week 2.

<table>
<thead>
<tr>
<th>CODE ID</th>
<th>AT</th>
<th>AS</th>
<th>SB</th>
<th>AP</th>
<th>AD</th>
<th>JS</th>
<th>JK</th>
<th>GC</th>
<th>BL</th>
<th>BR</th>
<th>KG</th>
<th>JF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem Type (A or B)</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>Similarity Rating (A/B vs. C)</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2.5</td>
<td>2.5</td>
<td>3.5</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Usefulness Rating</td>
<td>4</td>
<td>3</td>
<td>3.5</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td>Similarity Rating (A vs. B)</td>
<td>-</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Usefulness Rating** – Problem C was rated on how useful it was for assisting in the solution of the unsolved problems. Again, there was no observable difference in rating between students assigned problem A and problem B. On average, students rated problem C as 3.6 out of 5. Problem C was rated much higher in ‘usefulness’ than in ‘similarity’ with problems A and B.

**Problem Posing** – Students were unable to begin the problem posing task prior to the end of the interview.

6.3.4 *Focus group learning interview – Week 3*

**Methodology**

For week 3 of the focus group learning interviews, problems A, B, and C covered work-kinetic energy theorem. Figure 6.8 below shows the problems created for the interview. All problems may be viewed in full in Appendix D-2. Students were able to complete all tasks prior to the end of the session. One student from the cohort was missing from this week for family reasons.
**Figure 6.8 Problems used during FOGLI 3.**

**Problem C**  
Joshua pushes a 1.85 kg box along a flat horizontal table applying an average force of 39.0 N. The box starts at rest and reaches a velocity of 12.0 m/s. Neglecting friction, how far did Joshua push the box?

We may express the work done by Joshua on the box in terms of two quantities:

\[ W_{\text{net}} = \vec{F}_{\text{net}} \cdot \Delta x \]  

i. The force applied on the box by Joshua.  
ii. The distance covered while Joshua applied the force.

**Problem A**  
A 0.10 kg arrow is fired from a bow. The bow is pulled back a distance of 0.8 m so that the arrow is released with a speed of 50 m/s as it leaves the bow. The arrow travels 25.0 m before hitting its target. What is the average force exerted on the arrow by the bowstring?

**Problem B**  
A Yankees batter hits a 0.14 kg baseball sending it off into left field, 40 m away from the batter’s box. The baseball lands in a Royals fielder’s glove, exerting an average force of 300 N, moving the glove backward 0.25 m before coming to rest. What is the speed of the ball just before it is caught?

**Results for Week 3**

**Principles** – Students initially struggled with determining the principle underlying the problems for this week. Three of 11 students stated that the work-kinetic energy principle was necessary for their problem. Two students cited Newtonian mechanics as the primary principle.

Student JS “we need to find the speed of the ball and by doing this I will use the force and the distance to determine the speed.”

One student cited conservation of energy.

Student GC “Change in potential plus change in kinetic is equal to change in potential plus change in kinetic.”

Interviewer “This is the equation you will use?”

Student GC “Yeah. I think. I don’t know.”
The rest of our cohort described the problem statement in terms of what was asked for and what information was given.

Student AT “Looks like we're looking for the tension of the string given the speed and mass of the arrow and the distance pulled back…perhaps?”

Student AD “Well using the velocity of the arrow and it's mass and the distance it travels, we find the amount of force.”

Prior to this week, the moderators assumed this task was easier for our new cohort because they were given the example problem C right from the start of week 1. This assumption was clearly inaccurate. Students were given example problem C for this week prior to being asked to describe the principles involved in their own problem. Yet, students were clearly having trouble describing the principle prior to solving the problem.

**Text Editing** – All but one student selected the correct distance for their irrelevant information. Five of six students given problem A selected the 25 m travelling distance as irrelevant because the work done is before the arrow moves the 25 m.

Student AT “The target could be 2” or 1 mile away, doesn’t matter…unless we know the final velocity at the end and have resistance or something.”

Five of five students given problem B selected the 40 m as irrelevant because the work is done after the ball has travelled the 40 m distance.

Student JK “40 m is irrelevant because we only care how far the glove bends back.”

One student in the cohort, given problem A, was unsure and decided not to answer.

**Similarities and Differences** – The mean for similarity ratings were all above 2.5 out of 5. Students given problem A rated the problem a 2.92 out of 5 in similarity with problem C. Students given problem B rated the problem a 3.26 out of 5 in similarity with problem C. When asked to describe the similarity and differences between their unsolved problem and problem C, seven of the eleven students stated that the problems were primarily similar because they ‘use most of the same equations.’ The other four student participants stated that the problems shared the same principles and types of quantities.

Student BR “Both use work energy principle and both equate change in kinetic to net work scalar.”
Student GC “Same equation and setup. Neither problem took into account how
the normal and frictional forces were not involved.”

Student KG “Friction made no difference, but [both problems] used the same
equation and have similar setups.”

Students were also asked to describe the differences between their unsolved problem and
the example problem C. Students described the surface feature differences between
problems, the differences in quantities being solved for, and irrelevant information
contained only in problems A or B, not C. Interestingly, when students were asked to
rank these differences according to their importance to solving the problem, students
determined that the irrelevant information was the primary difference and the differences
in the quantities being solved for ranked second in importance.

Student AP “There is extra info in problem A; no extra info in problem C.
Problem A solved for force, but problem C solved for distance”

Students also rated the similarity between problems A and B. The mean rating for week
3 was 2.91. Students again described the similarities and differences between the two
problems. Two of 11 students stated that the primary similarity was the work-energy
principle. Nine of 11 students stated that the primary similarity was the equation used.
Three of these 11 students also cited the irrelevant information contained in the problems
was another, less important similarity. One student also noted that both problems were
projectile problems, though this student ranked this similarity as third in importance:

Student SB “Both problems use work energy principle, both try to throw you off
using different distances, and both work with projectiles.”

Table 6.3 Similarity and usability ratings for week 3.

<table>
<thead>
<tr>
<th>CODE ID</th>
<th>AT</th>
<th>SB</th>
<th>AP</th>
<th>AD</th>
<th>JS</th>
<th>JK</th>
<th>GC</th>
<th>BL</th>
<th>BR</th>
<th>KG</th>
<th>JF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem Type (A or B)</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>Similarity Rating (A/B vs. C)</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3.5</td>
<td>3</td>
<td>3.5</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3.8</td>
</tr>
<tr>
<td>Usefulness Rating</td>
<td>3</td>
<td>3.5</td>
<td>4</td>
<td>3.5</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Similarity Rating (A vs. B)</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2.5</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3.5</td>
</tr>
</tbody>
</table>

**Usefulness Rating** – Problem C was rated on how useful it was for assisting in the
solution of the unsolved problems. There was no observable difference in rating between
students assigned problem A and problem B. On average, students rated problem C as
3.27 out of 5. Problem C was rated equally as high in ‘usefulness’ and ‘similarity’ with problems A and B.

**Problem Posing** – All of the five groups of students participating in the week 3 focus group learning interview were able to complete the problem posing task and pass on their problems to other student pairs. Figure 6.9 below shows an example of a problem posed during week 3.

Of the five problems created, four were underspecified. All four problems were missing information pertaining to the distance work was done on an object. These three problems included distances, but they were distances travelled without an external force acting on them. Thus, it would be the wrong distance to use for a net work calculation. Only one pair of students were able to determine that the problem they were given to solve was underspecified. The other three groups given underspecified problems did not realize the distance given was incorrect.

Interestingly, three of these five problems involved football fields, and two of these three involved quarterbacks’ throwing footballs. The other two problems involved an arrow being shot by a bow and a cannonball being shot through the air and hitting a chicken. With exception to the bow and arrow problem, the student groups chose surface features that were different from the three problems presented in the interview. Two of these five problems were quite unique with respect to the selection of surface features. The third football problem involved a grandmother being shot through field goal posts, and the cannon problem involved a chicken.
6.3.5 Focus group learning interview – Week 4

Methodology

For week 4 of the focus group learning interviews, problems A, B, and C covered momentum and inelastic collision. Figure 6.10 below shows the problems created for the interview. All problems may be viewed in full in Appendix D-2. Students were able to complete all tasks prior to the end of the session.

Figure 6.10 Problems used during FOGLI 4.

Problem C

Astronauts are throwing around two wads of putty in a zero gravity environment when the two wads collide and stick to one another. Just before the collision, one wad, of mass 3.0 kg, is moving at a 40° degree angle downward from the vertical at 20 m/s. The other wad, of mass 2.0 kg, is moving upward at 12 m/s. Determine the velocity (direction and magnitude) of the two wads of putty after the collision.

SOLUTION:
We use the law of conservation of momentum (which states that the total momentum of the system is conserved when no external force acts)
Problem A
A 1200 kg semi-trailer truck (which includes a 400kg towing engine attached to a 800 kg trailer) travels on a straight line highway at the constant speed of 20.0 m/s while a film crew takes video for a new movie. After 15.0 seconds, a small explosion is set off at the hinge between the 200 kg towing engine and the trailer. The two parts separate, and the velocity of the trailer is 5.00 m/s in the direction opposite to the direction of the towing engine. What is the velocity of the towing engine after separation?

Problem B
A 35,000 kg railroad freight car moving at 2.0 m/s is initially 500 m behind a 25,000 kg car moving at 1.0 m/s in the same direction. After 1000 seconds, the cars make contact and stick together. Determine the speed of the cars after the collision.

Results for Week 4
Principles All twelve students described the central principle involved in the problem as momentum. Three of twelve students mentioned conservation of momentum specifically, and another four students stated that their problem included inelastic collisions. Student KG also asked for the moderator to explain why they were asked for the central principles, and how they were similar or different from the concepts.

Student KG “You see…I really don’t know what central principles mean…I assume momentum. What do you mean by central principles? Are they different from concepts?”

Moderator “Well, we can assume they are dependent on some overarching concept. For example, if we look at momentum, there are several kinds of momentum problems, elastic and inelastic collisions for example. When specifying principles, we want you to be as detailed as you can be about the physical phenomena.”

Student KG “Oh, gotcha.”

Text Editing – Sarcasm ran deep with this group during week 4. Multiple students, 7 of 12, wrote comments about how they really hoped there was not missing information. Prior to this interview, and at no time in the future, did students so adamantly express their dislike for this question. Students usually simply put a line through the question or wrote ‘NO’ next to the line.

Student KG “I hope not, cause I solved the problem anyways!”

Student AT “It would have been nice if we were told that the highway has no friction, but I guess you can’t have everything.”

228
Student JK “No, no, no…..should we just remove this question maybe?”

For the irrelevant information, students given problem A determined that the time given in the problem statement was unnecessary, and cited that the conservation of momentum does not require information about time.

Student BL “15.0 s is irrelevant. Time is not a component of momentum.”

Two of these seven students also stated that the 1200 kg was not necessary because the weight of the components were also supplied in the problem statement. The other students in the cohort agreed with these two students when this information was shared at the end of the interview.

Five of five students given problem B determined that the time and the 500 m distance were unnecessary. Four of the five students stated that these quantities were unnecessary because they were not apart of the equation. One of the five, student SB, stated they were not necessary because “There's no acceleration so this doesn't change anything.”

**Similarities and Differences** – Students rated problem A low in similarity as compared with problem C. Problem A, on average, rated 1.79 out of 5. Problem B was moderately similar with an average rating of 2.34 out of 5. All 12 students noted the primary similarity between problems was the principle, or inelastic collision.

Student AP “Both problems used the equation for conservation of momentum, total initial momentum + total final momentum, and they were both inelastic.”

Student JF “Your finding pi and pf [momentum initial and momentum final], you have a initial mass for both problems and velocity, and in the final momentum they both stick together. I would say that’s the more important of the three.”

Interviewer “Why would you say that the sticking together was most important?”

Student JF “Because it makes a difference in how the pi and pf [initial and final momentum] are set up in the equation, and the mass is really not that important of a similarity. Most problems have mass in them.”

Three students given problem B also noted that the problems both had movement in the x-direction.
Student AD “Both used the same x direction equation for momentum, Both use inelastic momentum”

Students who were given problem A noted that the problems were different because the inelastic collision ‘was reversed’ from one another. For problem A, the two objects were stuck together, then came apart; for problem C, the two objects were apart, then stuck together.

Five of seven students given problem A also noted that the problem A was on a single axis while problem C split off objects at an angle. These five students cited the additional geometry and mathematics as a primary difference between the problems. Students given problem B also cited the angle in problem C as a primary difference. Problem B also fixed movement to one axis.

Student BR “We didn’t have to find components…for my problem, the object experienced an explosion instead of collision, and only the x direction was involved.”

Student AP “Problem C was much more complex because it had an x- and y- axis. Problem B was easy and only had 1 axis, but problem C had to determine what was the sine and cos [cosine] for angles.”

When asked to compare similarity between problem A and problem B, students rated problem A and B, on average, 3.09 out of 5 in similarity. Problem A and B were rated much more similar to one another than when each individually unsolved problem was compared with the example problem C. Student recognized that problem A and B are primarily similar because of the inelastic conservation of momentum. Students also recognized that problem A and B are both constrained to move along one axis.

Student JS “same equation, both involved cars in motion, and both are on the x-axis.”

All students determined that the primary difference between the two problems was whether the interaction between the objects caused them to split apart or come together.

Student BL “One is a joining of two masses, the other is a separation. Problem A also involves a change of direction, but it’s still moving along the same axis.”
Table 6.4 Similarity and usability ratings for week 4.

<table>
<thead>
<tr>
<th>CODE ID</th>
<th>AT</th>
<th>AS</th>
<th>SB</th>
<th>AP</th>
<th>AD</th>
<th>JS</th>
<th>JK</th>
<th>GC</th>
<th>BL</th>
<th>BR</th>
<th>KG</th>
<th>JF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem Type (A or B)</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>Similarity Rating (A/B vs. C)</td>
<td>2</td>
<td>2</td>
<td>2.7</td>
<td>2</td>
<td>3</td>
<td>3.5</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Usefulness Rating</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3.4</td>
<td>4.8</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Similarity Rating (A vs. B)</td>
<td>3</td>
<td>3</td>
<td>3.5</td>
<td>2</td>
<td>4.5</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

Usefulness Rating – Problem C was rated on how useful it was for assisting in the solution of the unsolved problems. There was no observable difference in rating between students’ assigned problem A and problem B. On average, students rated problem C as 2.85 out of 5. Students’ ratings varied for this problem, and the variance was seemingly independent of which problem students were asked to solve. Problem C was rated equally high in ‘usefulness’ than in ‘similarity’ with problems A and B, particularly higher than problem A.

Problem Posing – All of the five student groups were able to complete the problem posing task and solve another student pair problem. All of the five problems were solvable. All student groups perceived the problems as solvable. One of the five problems involved a ball of clay which is a surface feature similar to problem C. Two of these problems involved vehicles colliding (a Volkswagon Beetle collides with a semi-trailer or a Ford Mustang collides with a Lincoln Town Car), similar in surface features to problem A. Another problem involved three velco balls being thrown at a butterfly and colliding with one another while squishing the insect. The final problem involved two bullets of different masses and speeds being fired at one another and colliding.
6.3.6 Focus group learning interview – Week 5

Methodology

For week 5 of the focus group learning interviews, problems A, B, and C covered static fluids and pressure. Figure 6.12 below shows the problems created for the interview. All problems may be viewed in full in Appendix D-2. Students were able to complete all tasks prior to the end of the session.
**Problem C**

In working out his principle, Pascal placed a long, thin tube of radius \( r = 0.0030 \text{ m} \) vertically into a barrel of radius \( R = 0.21 \text{ m} \) and height 0.80 m. He found that when the barrel was filled with water, and the tube filled to a height of 12 meters, the barrel burst. Calculate:

(a) The mass of the water in the tube.
(b) What is the pressure exerted by the water at the bottom of the tube just before rupture.
(c) The net force exerted by the water in the barrel on the lid just before rupture.

**Problem A**

The maximum gauge pressure in a hydraulic lift is \( 1.722 \times 10^{8} \text{ Pa} \). What is the largest mass of the vehicle it can lift if the radius of the output line is 0.14 m length is 1.2 m?

**Problem B**

You have a swimming pool of length 22.0 m and width 8.5 m. The uniform depth of the pool is 2.0 m. What is the pressure due to the water against the side of the pool near the bottom?
Results for Week 5

Principles – Students were able to identify the primary principles involved at pressure and force in fluids. Three of the 12 students also mentioned that Pascal’s principle was the primary principle involved.

Student AP “The principle of this problem is to determine what mass the hydraulic lift can raise, and it depends on the output line. It is related to pressure and force.”

Student GC “Pascal’s principle is needed, we need to solve for the pressure exerted by the H2O on the wall near the bottom.”

Text Editing – Students given problem A determined that the 1.2 m given in the problem statement was not necessary because it did not affect ‘the system.’

Student JF “1.2 m is irrelevant because it doesn’t matter the length of something not in the system.”

Student AS “1.2m. It’s not needed to get the mass, and we already have the radius. It doesn’t have anything to do with the part that we focus on.”

Students given problem B determined that the 22.0 m and 8.0 m was unnecessary information because there is no need to calculate the volume.

Student AD “22.0 m and 8.5 m is not needed. The volume of the pool is not asked for, if it was asked for we could use this info, but were not asked about that.”

Similarities and Differences – Students rated problem A moderately similar to problem C. Problem A, on average, rated 2.42 out of 5. Problem B was also rated moderately similar to Problem C with an average rating of 2.41 out of 5.

All 12 of 12 students recognized that a pressure was being exerted on something in all the problems. This feature was cited as the primary similarity between problems A/B and C. Four students also stated that problems A/B and C were similar because they used water as the interacting liquid.

Student JS “There is a force exerting on something and they will a similar equation because of that. I don’t know, they also both have water.”
Students given problem A stated that the problems primary difference was that the unsolved problem A did not find the force, but used a given weight. Two students also specifically stated that there was no use for the density of water in their problem.

Student BL “No use of water density, and fluid force versus gravity.”

Student AS “Used the weight of the car instead of finding the force.”

Students given problem B stated that the problems were different because one solved for a force while the other solved for pressure. One student given problem B also stated that their problem contained more irrelevant information.

Student BR “Only pressure was calculated, we didn’t have to calculate mass or net force. We had more irrelevant information.”

When asked to rate the similarity between problem A and problem B, students rated problem A and B, on average, 1.58 out of 5 in similarity. Problem A and B were rated much less similar to one another than when each of them individually was compared with the example problem C. All 12 students stated that the problem were similar because they used pressure.

Students agreed that there were many more differences than similarities. Student BR below summarized the list of differences students collaboratively talked about:

Student BR “There are so many more differences: different calculations, the gauge pressure principle applied in problem A, the volume was needed in problem A, no density required in problem A.”

Table 6.5 Similarity and usability ratings for week 5.

<table>
<thead>
<tr>
<th>CODE ID</th>
<th>AT</th>
<th>AS</th>
<th>SB</th>
<th>AP</th>
<th>AD</th>
<th>JS</th>
<th>JK</th>
<th>GC</th>
<th>BL</th>
<th>BR</th>
<th>KG</th>
<th>JF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem Type (A or B)</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Similarity Rating (A/B vs. C)</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>1.5</td>
<td>2.5</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Usefulness Rating</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2.7</td>
<td>3.5</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Similarity Rating (A vs. B)</td>
<td>2</td>
<td>1.5</td>
<td>1</td>
<td>1.5</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

**Usefulness Rating** – Problem C was rated on how useful it was for assisting in the solution of the unsolved problems. There was no observable difference in rating between students’ assigned problem A and problem B. On average, students rated problem C as 2.85 out of 5. Problem C was rated equally high in ‘usefulness’ and ‘similarity’ with problems A and B.
Problem Posing – All of the six student groups were able to complete the problem posing task and solve the problems created by another group of students. All of the five problems were solvable, though one problem was given a caveat after the problem was already handed to another student pair. All student pairs perceived the problems as solvable. There were many unique scenarios given for these problems: a water main break, a grandmother lifting a medicine ball filled with water, a giraffe extending his head downward, a fish pulling a man into a lake and a KY Jelly wrestling match. The KY Jelly problem was originally underspecified, as the person was not submerged and a mass was not given for the person. The student pair that created the problem realized this prior to talking with the student pair that attempted to solve the problem, and corrected the problem by walking over and telling the student pair to look for the pressure on the side of the person’s flip flops, not the bottom of the person’s flip flops. This was an acceptable change to make the problem solvable. See Figure 6.13 below for an example of a problem posed. The white block is placed in the figure to remove names of participants.
Figure 6.13 Example of a problem posed given in week 5.

The normal pressure surrounding a giraffe’s brain is 240 mmHg systolic when standing. However, when the giraffe’s head is lowered to get a drink ($\Delta 4m$), the cardiovascular system must make significant changes in blood pressure to keep the giraffe from developing brain damage due to excess pressure. Assuming blood density similar to human blood, what is the potential change in pericranial blood pressure per standing to drinking height?

$$\delta P = \rho g h$$

$$\delta P = (1050 \text{ kg/m}^3)(9.81 \text{ m/s}^2)(4 \text{ m}) = 4.2 \times 10^4 \text{ Pa}$$
6.3.7 Focus group learning interview – Week 6

Methodology

For week 6 of the focus group learning interviews, problems A, B, and C covered simple harmonic motion. Figure 6.14 below shows the problems created for the interview. All problems may be viewed in full in Appendix D-2. Students were able to complete all tasks prior to the end of the session.

Figure 6.14 Problems used during FOGLI 6.

Problem C
A spring is mounted on a frictionless 60° incline with a spring stiffness constant of 350 N/m. The spring vibrates with an amplitude of 0.20 m when a 0.30 kg mass hangs from it. Assume the incline is 2 m long, and the mass has a positive velocity when passing through equilibrium at t=0.
(a) What is the frequency of the motion?
(b) What is the equation of motion?
(c) What is the maximum velocity obtained by the mass?
(d) What is the total energy?

Problem A
A 0.20 kg mass is attached to the end of a horizontal spring bolted against a wall. When the spring is fully compressed, the distance from the wall to the midpoint of the mass is 2.1 m. Assuming there is no friction between the mass and the floor, the mass is released and the system begins to oscillate. The equation that describes the motion of this system as a function of time is
\[ x = 0.7 \cos(2\pi(3)t) \]

Problem B
A vertical spring with spring stiffness constant 305 N/m vibrates with an amplitude of 0.20 m when 0.60 kg mass is set on top of it. The mass passes through the equilibrium point (y=0) with a positive velocity at t=0. Assuming the mass is 0.40 m away from the ground at full compression,
(a) What is the frequency of the motion?
(b) What equation describes this motion as a function of time?
(c) What is the maximum velocity obtained by the mass?
(d) What is the total energy?
Results for Week 6

**Principles** – Students were able to identify the principle as simple harmonic motion. Four of 12 students referred to the principle as ‘harmonic motion.’ Three of 12 students referred to the principle as ‘spring oscillations,’ and five of 12 students referred to the principle as ‘simple harmonic motion of a spring.’

Student AP “The central principles of the problem is to use it's mass, amplitude and a given constant to solve a variety of equations that are in regards to energy, velocity and frequency of springs.”

Interviewer “So if you had to describe this principle in a short few words, how might you do so?”

Student AP “It’s spring motion principles I guess.”

**Text Editing** – Students given problem A determined that ‘2.1 m’ was unnecessary. Four out of six students cited that the ‘2.1’ was not necessary because one does not need to know the start position, only the displacement. Two of the six students stated that the 2.1 was not necessary because the equation used did not include that piece of information.

Student JK “2.1 m, it doesn't matter how far it is to start, only the displacement matters.”

Student SB “2.1 m”

Interviewer “Why is that not necessary?”

Student SB “It isn't used in any equations.”

Students given problem B determined that ‘0.40 m’ was not necessary because it was not apart of the equations.

Student AD “0.40 m was unnecessary because I don’t need to use for anything.”

Student JF “0.40 m is unnecessary. I can’t think of what I would plug it into.”

**Similarities and Differences** – Students rated problem A high in similarity as compared with problem C. Problem A, on average, rated 3.10 out of 5. Problem B was also rated high with an average rating of 3.58 out of 5. The primary similarity between problems A/B and C was that they were simple harmonic motion problems. Students also stated that the problems were similar because they used similar equations and contained springs.

Student AP “Used all the same equations, both used a spring on some axis. These are basic simple harmonic problems.”
Student BL “I would say simple harmonic motion…as the main similarity…both were springs moving after a weight was introduced, and both used similar equations…no wait, they used the same equations [shakes head in agreement with himself].”

Students given problem A and B stated that the primary difference between their unsolved problem and problem C was the angle given in problem C.

Student AD “Problem B was on a vertical plane and problem c was on a 60° incline.”

Student JF “The angle of problem c was at a slope with a angle, our problem b was straight up.”

Four of six students given problem A stated that they needed to calculate ‘ω’ in their problem but not in problem C. One student given problem B also stated that problem C contained more useless information than their problem.

When asked to compare similarity between problem A and problem B, students rated problem A and B, on average, 3.18 out of 5 in similarity. Problem A, B, and C all rated fairly high in similarity with one another.

Again, students stated that the problems were similar because they involved simple harmonic motion. Student felt the problems were different from one another because they asked for and gave different quantities. Students also stated that the problems were different because one was vertical while the other was horizontal.

Student AP “problem A did not have to solve part b that problem B did, problem B was vertical; Problem A was horizontal.”

Student BR “Problem A calculated k, problem B’s equation was calculated as a system as a function of time. They were different representations of the same problem. Problem A horizontal motion and problem b vertical motion.”

Table 6.6 Similarity and usability ratings for week 6.

<table>
<thead>
<tr>
<th>CODE ID</th>
<th>AT</th>
<th>AS</th>
<th>SB</th>
<th>AP</th>
<th>AD</th>
<th>JS</th>
<th>JK</th>
<th>GC</th>
<th>BL</th>
<th>BR</th>
<th>KG</th>
<th>JF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem Type (A or B)</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>Similarity Rating (A/B vs. C)</td>
<td>3</td>
<td>-</td>
<td>3.5</td>
<td>3.5</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3.5</td>
<td>2</td>
<td>4</td>
<td>3.5</td>
</tr>
<tr>
<td>Usefulness Rating</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Similarity Rating (A vs. B)</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3.4</td>
<td>3.5</td>
<td>3</td>
<td>3.5</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3.7</td>
</tr>
</tbody>
</table>
Usefulness Rating – Problem C was rated on how useful it was for assisting in the solution of the unsolved problems. There was an observable difference in rating between students’ assigned problem A and problem B. On average, students rated problem C as 3 out of 5 in usefulness as compared to problem A. On average, students rated problem C as 4 out of 5 in usefulness as compared to problem B. Problem C was rated equally high in ‘usefulness’ and ‘similarity’ with problems A and B.

Problem Posing – Four of five student pairs (two groups of three) were able to complete the problem posing task and solve another student pair problem. One group ran out of time prior to creating their own problem. All four problems were solvable. All student pairs perceived the problems as solvable. The problems varied in scenario, though not as much as previous problems. The concept of simple harmonic motion is difficult to extend beyond the contexts of springs and pendulums, so this lack of variation is expected. The variations of posed problems were as follows: a guy bouncing on a pogo stick, a vertical spring with a person sitting on top, a dashboard bobble-body with a spring that compresses and expands, and a grandfather clock with a pendulum swing.
6.3.8 Summary of Results

The focus group learning interviews were seamlessly implemented during this second implementation. The protocol for the interviews were finalized in the previous semester, and often, problems from the previous semester were used during the second semester implementation.

**Principles and Irrelevant information.** Students were asked to identify the principles involved within the problems while they worked on a solution. Students were often able to identify the principle involved and describe the situation thoroughly. Students were also able to identify the irrelevant information given in the problem and provide sufficient explanation for why they chose the irrelevant information. It is likely
that the example problem C assisted students with the task of choosing appropriate concepts and irrelevant information, but there is some evidence (from week 3 primarily) that supports the idea that students are not entirely dependent on the example problem for these tasks. Our students were very candid while explaining their answers or observations. It would be possible that the process of utilizing an analogy may be more sub-conscious, though it would be expected that at least one person would cite the example (instead of an equation which is often the case) as evidence to their claim.

Unlike the previous semester, students were capable of identifying the irrelevant information without solving the unsolved problem first. It is unknown exactly as to why this may be true, but it might suggest that the ordered stepwise procedural scaffolding provided to the students during these interviews was more effectively enforced by the moderators. It is possible that the large number of initial modifications made to the worksheets during the previous semesters’ group learning interviews might have eliminated certain internally imposed obligations to follow directions.

**Similarities and Differences.** Students were very open about discussing similarities and differences. Similarities included surface features of a problem, but the primary similarities chosen by students were entirely deep-structure elements of the problems. Differences were often surface feature related, but it should be noted that there is a deep emphasis on mathematical differences between problems. This is seen very obviously in week four, when students cite the primary difference between the unsolved problems and the example problem C as being the number of axial components necessary to be calculated. The problem similarity ratings often varied from problem to problem. Rarely were problems rated very low in similarity with one another. Rarely were problems rated very high in similarity with one another. This feature was not included in the previous semester protocol, but was added to get a better quantitative feel for students’ perception of similarity versus usability of example problems.

**Usefulness of example problem.** The usefulness of example C varied over the semester, though it was often higher than 2.5 out of 5. The usefulness rating and the similarity ratings were not comparative from week to week, as sometimes students would rate the example problem as very useful, but also rate the problem as only moderately similar to their unsolved problem. This difference in ratings was not unexpected.
Surface feature differences, and additional mathematics, like simple conversions or breaking down angles, may not be as important with regard to choosing a useful solved example, but they are differences.

**Problem posing tasks.** Students were asked to pose problems that incorporated elements from the three problems utilized during each interview. Students were unable to complete this task for the first two weeks due to lack of time. For the third week, all students in the cohort created problems, but three of the five problems were underspecified. For the remaining weeks, students were able to create problems that were sufficiently specified and many of them could be classified as unique scenarios. There were no noteworthy changes in the way students created problems from week to week. Note instead that students’ selection of surface features was somewhat dependent on how similar surface features were within the problems A, B, and C given for the interview. If problems A, B, and C all contained springs, then the problems posed would also all contain springs. This quality of ‘uniqueness of surface features’ is also a trend that was observed in the previous semester.

### 6.4 Individual Interviews

The problems used for our final individual interview in our previous semester were reused for the individual interview conducted for this study. All problems developed shared an overarching concept of simple harmonic motion. Four distinct problem scenarios were chosen, (i.e., a baby in a bouncy chair, a rock falling on a vertical spring, etc) each incorporated one of two primary conditions (i.e. no elastic potential energy/elastic potential energy, swinging pendulum motion/oscillating spring) governed by the overarching concept.

Each distinct problem scenario was used to create a set of three problems of varying deep-structure similarities. The first and second problems would be solvable using the same equations, but with different given and unknown information provided in the problem statements. The third problem was conceptually similar but required a different physical principle (e.g., simple harmonic motion as a function of time) than the first and second problems. The third problem would remain surface feature similar to the first and second problem, and the primary concept remained the same, but the different
principle would critically alter any equation(s) used for problem 1 and 2. Thus, the third problem was structurally different as compared with problem 1 and 2.

Because of this creation of problem sets, there always existed two separate sets of problems that could be paired in such a way as to create all variances in surface feature similar/different and deep-structure similar/different. Figure 6.16 represents how problem sets were formed. This is similar to the diagram seen in chapter 5. The representation only includes 2 scenarios, though our actual problem sets had four scenarios. Figure 6.17 represents how different pair types might be formed using the problems. Again, a similar figure was shown previously in chapter 5.

**Figure 6.16 Diagram representing two scenario sets.**

C: Concept (e.g., SHM)
P: Principle (e.g., function of time includes trigonometry, \( x = A \))
QS: Quantity Solved for
The individual interviews required students to perform two separate tasks all involving the problems described above. Problems used for the ‘usability task’ were not the same problem set as used in the ‘similarity ratings task’. Students participants were randomly split into two groups such that problems given for the usability task for group 1 were the same problems used for the similarity ratings task for group 2 and vice versa.

### 6.4.1 Similarity Rating Task

During the individual interviews students were asked to rate the similarities between contrasting problems of varying deep structure and surface feature similarities.
We presented students with pairs of problems and asked them to rate the similarity of each pair on a five-point Likert scale. Each student was presented with eight pairs of problems. The problem pairs of were constructed from problems that had facial similarities/differences and principle similarities/differences. The term facial similarity/difference corresponds to surface feature similarity/difference, while the term principle similarity/difference corresponds to deep structure similarity/difference.

All four combinations of facial/principle similarities/differences were created. These are labeled problem pair types A, B, C, and D as defined in the Table 1 below.

Table 6.7 Problem pairs for the similarity rating task.

<table>
<thead>
<tr>
<th>Principle Similarity (PS)</th>
<th>Facial Similarity (FS)</th>
<th>Facial Difference (FD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principle Difference (PD)</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>D</td>
</tr>
</tbody>
</table>

Students were presented the problem pairs in order A, A, B, B, C, C, and D, D. Students were not allowed to backtrack and change their similarity rating for any pair until the end of the sequence when they were given the opportunity to review their ratings for all pairs and decide whether they wanted to revise any of the similarity ratings.

Four graduate-level faculty were also given the same similarity rating task. These faculty were chosen because they had multiple years of experience teaching introductory-level algebra-based and calculus-based physics. They also had an interest in physics education research and used some level of research-based teaching methodology in their own classrooms. None of the four faculty were physics education researchers. The four faculty ratings were used to determine whether the actual faculty ratings aligned with how the researchers expected the faculty to rate the problem pairs.

6.4.2 Results – Similarity Rating Task

In our previous phase, Phase II, we asked students to do this same task with the same problems in their second interview. In Phase II, we found that the differences between problem types A and B, and B and C were no longer statistically significant as compared with the first interview. Students rated pairs A and B at about the same level of similarity. These data were consistent with the notion that students recognize that the
problems in pair B have principle similarities that overpower their facial differences to the extent that they rate pair B almost the same way as they rate pair A. The Figure 6.18 below shows the mean ratings from Phase III’s individual interview.

**Figure 6.18 Similarity ratings for the individual interview in phase III.**

For our similarity rating task during this interview, students rated the same four types of pairs at the end of the semester following the focus group learning interview treatment. Students rated pair types A, B and C all at about the same level of similarity, remaining consistent with the previous semester. In hindsight, it would have been useful to compare interview 1 data from our last semester with our new set of students, but as discussed earlier, there was no time in the semester to complete the first interview.

Figure 6.19 below shows the mean ratings for each problem type for the individual interview conducted in Phase III and the individual interview 2 conducted the previous semester.
Students were also asked to discuss their ratings with the interviewer. As in Phase II, students definitively noted that the problems given in all pairs shared the same concept.

Student BL “Well, all of these problems are alike. They are all problems involving simple harmonic motion, and I could rate them all very close. In fact I did!…..this pair and this pair (last two pairs, type D) are still similar to the rest, but they require just a little more work, so they got a slightly lower number than the rest.”

Four faculty were also asked to rate the problem pairs. Due to the small number of faculty participants, a statistical analysis is not appropriate. The faculty ratings were only used to determine whether the actual faculty ratings aligned with how the researchers expected the faculty to rate the problem pairs. Previous research (Chi, Feltovich et al., 1981; Chi, Siler et al., 2001) suggests that expert physics problem solvers emphasize physical principles over facial features. It would be expected then that problem pair type A would rate highest with both principle and facial similarity, problem pair type B would rate second highest with principle similarity and facial differences, problem pair type C would rate third highest (or second lowest) with facial similarity and principle differences, and finally problem pair type D would rate lowest with principle and facial differences. After faculty ratings were averaged, it was apparent that the
faculty did in fact rate the problem pairs as expected. Their average ratings were plotted over the top of the end interview ratings from Phase II and Phase III as seen below in Figure 6.20. It may be noted that Phase II and Phase III students’ ratings for three of the four problems are similar to the faculty' ratings by the end of the semester. Problem type C is most different. Students rate type C problem pairs higher than type A and B problem pairs, while faculty rate type C problem pairs lower than type A and B problem pairs.

Figure 6.20 Faculty and end interview ratings from Phase II and Phase III.

6.4.3 Usability Task

For the second task given during the individual interview, students were asked to view six problems and predict which of the six problems would be most and least useful as a solved example to enable them to solve a challenging problem that was provided. The problems were taken from two of the four problem scenarios. The two problem scenarios were chosen such that they did not have corresponding primary conditions (spring vs. pendulum). Exactly half of the students were assigned the problems that were a part of set 1 (the first two problem scenarios) and the other half were assigned the problems that were a part of set 2. Students were presented with a challenging problem, seen in Figure 6.21 below. The challenging problem remained the same for both sets.
The challenging problem given for this interview was identical to the problem given during interview 2 of Phase II.

Figure 6.21 Challenging problem used in the usability task.

![Diagram of a simple pendulum and spring system with labels: 0.80 m, 0.30 kg, 0.50 kg, and k=7.70x10^9 N/m.]

The length of a simple pendulum is 0.80 m, the pendulum bob has a mass of 0.30 kg, and it is released at an angle of 12° to the vertical. At the lowest point of the swing, the bob strikes a 0.60 kg block attached to a fixed horizontal spring whose spring stiffness constant is 7.70x10^9 N/m. The block is set into vibration.

(a) What is the speed of the bob before it strikes the block?
(b) What is the resulting amplitude of the block after being struck by the bob?
(c) At what angle would we have to start the simple pendulum if our block is set into vibration with an amplitude of 0.20 m?

6.4.4 Results – Usability Task

Students given set 1 chose a wide variety of problems for most and least useful. Students chose problems 5, 6, 10 and 12 as most useful, and problems 5, 6, 11, and 12 as least useful. Problems chosen from set 1 as most and least useful are listed in Table 6.8 below.
Table 6.8 Usability problems selected by students for set 1 as most and least useful.

<table>
<thead>
<tr>
<th>Problem No.</th>
<th>Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td><img src="image" alt="Diagram" /> A common device for entertaining a toddler is a jump seat that hangs from the ceiling. Assume that jump seats’ elastic cords act like a large spring. Once the 15 kg toddler is placed in the seat, a caregiver pulls down the jump seat and releases, allowing the toddler to oscillate up and down taking 3 seconds to complete one oscillation. The toddler never moves faster than 0.20 m/s. Find the maximum distance the toddler moves away from equilibrium and spring stiffness constant of the elastic chords.</td>
</tr>
<tr>
<td>6 (prob. 3 of scenario set)</td>
<td><img src="image" alt="Diagram" /> A common device for entertaining a toddler is a jump seat that hangs from the ceiling. Assume that jump seats’ elastic cords act like a large spring of stiffness constant 60 N/m. Once the 15 kg toddler is placed in the seat, a caregiver pulls down the jump seat an extra distance of 0.50 m from equilibrium and releases. Find the position and velocity of the toddler at ( t=4 ) seconds after she was released.</td>
</tr>
<tr>
<td>10</td>
<td><img src="image" alt="Diagram" /> A 2500 kg demolition ball swings from the end of a crane. The length of the swinging segment of cable is 20 m. The ball starts from rest 5.0 m above its lowest swinging point. Find the period of motion and the maximum velocity of the ball.</td>
</tr>
<tr>
<td>Problem No.</td>
<td>Problem</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
</tr>
<tr>
<td>11</td>
<td><img src="image1.png" alt="Diagram" /> A 2500 kg demolition ball swings from the end of a crane. The ball is pulled back from its vertical position and released so that it strikes the wall at the opposite end of its swing 0.7 seconds after it was released. The ball at the end of the cable acquires a maximum velocity of 10 m/s. Find the length of the swinging segment of the cable and the angle with respect to the vertical at which the ball was released.</td>
</tr>
<tr>
<td>12 (prob. 3 of scenario set)</td>
<td><img src="image2.png" alt="Diagram" /> A 2500 kg demolition ball swings from the end of a crane. The length of the swinging segment of cable is 20 m. The ball starts from rest 5.0 m above its lowest swinging point. Find the velocity of the ball and the angle it makes with the vertical at t=0.20 seconds after it was released.</td>
</tr>
</tbody>
</table>

Three of six students chose the pendulum problems as most useful, while these same three chose the spring problems as least useful. Similarly the other three students chose the spring problems as most useful while choosing the pendulum problems as least useful.

Student GC “I don’t understand springs as well as I get pendulum problems. I think this one would be most useful.”

Interviewer “So, there are three problems involving springs, what was it about this one that made you choose this over all others?”

Student GC “Well I didn’t want to choose this one because it looked harder(problem 3 of the scenario set), so I just picked between these two (problem 1 and 2 of the scenario set)”
Table 6.9 Student selections of problems for the individual interview.

<table>
<thead>
<tr>
<th>Interview (SHM)</th>
<th>PENDULUM 1</th>
<th>PENDULUM 2</th>
<th>PENDULUM 3</th>
<th>SPRING 1</th>
<th>SPRING 2</th>
<th>SPRING 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOST Set 1</td>
<td>1</td>
<td></td>
<td>2*</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LEAST Set 1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2*</td>
<td></td>
</tr>
<tr>
<td>MOST Set 2</td>
<td>3</td>
<td>2</td>
<td></td>
<td>1*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEAST Set 2</td>
<td></td>
<td>1*</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
</tbody>
</table>

The table shown above represents students’ selections for most and least useful problems. The ‘Pendulum 1’ title represents the first problem of the scenario sets which contain a pendulum. Similarly ‘Spring 1’ title represents the first problem of the scenario sets which contain a spring. There are two problems, one from each interview set, that will coincide with each of these titles. The ‘*’ indicates that those data points are the same people within the set.

For all six students participating in the individual interview using set 1, their choice for the most and least useful problem was directly related to the perceived difference in difficulty between spring and pendulum harmonic motion problems. If a spring problem was perceived as more difficult, than a spring problem was chosen as most useful. Students selecting problems as most and least useful also perceived the third problem of each scenario set as more difficult compared to the other two problems of similar scenario. It is important to note that the third problem of every scenario set requires the use of a different equation, but the problem is not more conceptually or mathematically complex than the other problems. Three students chose problem 3 (principle different) problems because they were ‘more difficult to work out and could help with other problems.’ The other three students chose problem 2 because they were ‘having difficulty with those types of problems so looking at a simple example might help them with a more challenging problem.’

Students given set 2 chose problems 2, 8 and 7 as most useful while selecting problems 3 and 8 as least useful. See Table 6.10 below for problem descriptions. All but one student chose pendulum problems as most useful because they were ‘more difficult to understand than spring problems.’ These students also chose spring problems as least useful because they ‘understood spring problems better.’

254
Student KG “I like the spring ones.”

Student AD “The springs are easier than the pendulums, I think. I don’t like when we calculate angles.”

One outlying student also chose problems based upon their perceived difficulty, but his difficulty was with spring problems, not pendulum problems.

Student AS “I just don’t get these….these um…spring problems. I think I liked them before when we had a spring constant, but they are not same this time.”

Similar reasoning was used during the individual interview 2 of the previous semester, though it was not as prominent as an obvious emergent pattern. It would be interesting for future study to see whether this same perceived difficulty in reasoning is as prominent across many different physical concepts, or if simple harmonic motion is a unique outlier.

Table 6.10 Usability problems selected by students for set 2 as most and least useful.

<table>
<thead>
<tr>
<th>Problem No.</th>
<th>Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td><img src="k.png" alt="Diagram" /> A 0.40 kg mass is sitting on a 30° inclined frictionless surface and is attached to the end of a spring vibrating 3 times per second. The other end of the spring is bolted to the top of the 3.0 m high incline. The mass is stretched 0.30 m from equilibrium and released. The maximum velocity of the mass during oscillation is 0.55 m/s. Find the amplitude of the oscillation and the spring stiffness constant?</td>
</tr>
<tr>
<td>Problem No.</td>
<td>Problem</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
</tr>
<tr>
<td>3</td>
<td>A 0.40 kg mass is sitting on a 30° inclined frictionless surface and is attached to the end of a spring with a spring stiffness constant of 300 N/m. The other end of the spring is bolted to the top of the 3.0 m high incline. The mass is stretched 0.30 m from equilibrium and released. Find the position and velocity of the mass at t=3 seconds after it was released.</td>
</tr>
<tr>
<td>7</td>
<td>A pendulum consists of a 3.0 kg uniform disk fixed at the end of a 1.0 m long rod of negligible mass hanging from a frictionless pivot. The pendulum is tilted from its vertical position and released. The center of the disk is 0.30 m above its lowest swinging point prior to release. Find the time period of the pendulum and the maximum velocity of the disk.</td>
</tr>
<tr>
<td>8</td>
<td>A pendulum consists of a 3.0 kg uniform disk fixed at the end of a rod of negligible mass hanging from a frictionless pivot. The pendulum is tilted from its vertical position and released. The pendulum completes 4.0 oscillations per second, with the disk acquiring a maximum velocity of 2.0 m/s. Find the length of the rod and the angle with respect to the vertical at which the pendulum was released.</td>
</tr>
</tbody>
</table>
6.5 Non-Traditional Problems on Examinations

Data were also collected from five multiple choice examinations taken during the semester. Individual scores for each examination question were obtained by the primary course instructor. The last three problems on each examination were adaptations of text editing (Low and Over, 1990), physics jeopardy (Van Heuvelen and Maloney, 1999), and problem posing tasks (Mestre, 2002). The traditional examination problems assigned by the instructor, unlike the previous semester, were open-ended word problems. While the non-traditional tasks in their original form are open-ended, the problems included on the exams were in multiple-choice format. This was done such that they could be graded efficiently for a large number of students and also so that the data collected might be comparable across semesters. We acknowledge that the open-ended tasks can provide richer information about the students’ conceptual knowledge, but we could not process such vast amounts of open ended data efficiently. Problems created for the previous semester were reused for this study, but the examinations did not overlap physics content in the same way as the previous semester. In other words, examination 5 from phase II did not cover all of the same material as examination 5 from phase III. Examination 1 also fell early for this semester and was unevenly spaced as compared with the rest of the examinations. Problems included on the examination that were classified as ‘traditional’ were not simple ‘plug and chug’ problems like those seen in the previous semester. They often required the use of multiple equations and information given in the problem statement included unnecessary detail. There were also fewer problems overall compared to the Phase II. For the phase II semester, there were, on average, 35 traditional multiple choice problems per examination. For this phase, there were, on average, 11 traditional word problems per examination. In hindsight, it may have been appropriate for the investigator of this project to ask the instructor to explain their goals for the course. It is often assumed that the instructors aim to improve students’ conceptual understanding and problem solving ability, but the differences in importance placed on conceptual understanding versus problem solving techniques by the instructors in Phase II and III remains unknown. See Figure 6.22 below for the timeline with respect to the
examinations and the focus group learning interviews. See Appendix D for full non-traditional extra-credit problems for each examination.

Figure 6.22 Timeline for examinations and FOGLI's.

6.5.1 Text Editing

Text editing tasks involve presenting a student with a problem statement and then asking the student to identify the missing, irrelevant, and required information in the problem statement without first solving the problem. Low and Over (Low and Over, 1990) state that text-editing tasks can be a measure of schematic knowledge because they require an understanding of the deep structure of the problem. Because students are asked to complete the tasks without solving the problem, students need to know the interrelationships between various physical quantities, not in terms of equations, but at a conceptual level to be able to successfully complete the task. Figure 6.23 below shows an example of text editing used on one of the class exams.

Figure 6.23 Example of a multiple choice adaptation of a text editing task.

You are given a problem below.

A 2.0 kg mass initially 1.0 m above the ground is attached to thin cord that passes over a frictionless pulley to a second 3.0 kg mass which is initially 4.5 m above the ground. Both masses are initially at rest. Find the final velocity of the 3.0 kg mass right before it hits the ground.

In the problem statement above, specify which, if any, of the following quantities are not relevant for solving the problem.

(a) 2.0 kg mass  (b) 3.0 kg mass  (c) 4.5 meters  (d) 1.0 meters  (e) None of the above. You need all the information given to solve the problem.
6.5.2 Physics Jeopardy

Physics Jeopardy tasks were first developed by Van Heuvelen and Maloney (Van Heuvelen and Maloney, 1999). As the name indicates, these tasks require the students to work backward. Students are given a fragment of a solution to a problem and asked to identify the physical scenario that corresponds to the solution. The developers point out that these tasks require an effort to represent a physical process in a variety of ways. Because of these features, students are unable to use naïve problem solving strategies while solving Jeopardy problems.

Figure 6.24 below shows an example of our adaptation of a Jeopardy problem that provides students with a few steps of a projectile motion. Students are asked to determine what trajectory shown corresponds to the problem. This task requires students to relate information given in the mathematical and symbolic representation to a visual or pictorial representation.

Figure 6.24 Example of a multiple choice adaptation of physics Jeopardy task.

6.5.3 Problem Posing

Problem posing tasks were used by Mestre and others (Mestre, 2002) in the context of physics problems. In the tasks presented by Mestre, students were given a scenario, typically in the form of a picture and were asked to construct a problem around the scenario that was based on certain physical principles. Mestre points out that problem posing tasks are aimed at probing students’ understanding of concepts as well as
assessing whether they transfer their understanding to a new context. Our adaptation of this task is much more focused than Mestre’s original open-ended task. It presents students with the first part of a problem statement which clearly describes a physical scenario. Students are then asked to select from a list of choices, a question, which when added to the statement will create a solvable problem that requires the use of a set of given equations. Clearly, our adaptation differs from the original problem posing task designed by Mestre. First, this task clearly does have a unique correct answer. Second, it requires the knowledge of specific conceptual knowledge, represented in the form of equations. An example of our adaptation of a problem posing task is shown in Figure 6.25 below.

Figure 6.25 Example of a multiple choice adaptation of a problem posing task.

<table>
<thead>
<tr>
<th>You are given the starting statement of a problem below.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 500 kg cargo shipment, attached to a parachute, drops vertically out of a helicopter hovering 100 m above a large spring (k = 220,000 N/m). The cargo comes to rest when the spring compression is 0.50 m.</td>
</tr>
<tr>
<td>Which question, when added to the statement above, will make a solvable problem that requires ALL of the following equations to solve?</td>
</tr>
<tr>
<td>[ W = Fd \quad W = \Delta KE + \Delta PE \quad PE_{\text{spring}} = \frac{1}{2} kx^2 \quad PE_{\text{gravity}} = mgy \quad KE = \frac{1}{2} mv^2 ]</td>
</tr>
<tr>
<td>(a) What is the speed of the cargo just before striking the spring?</td>
</tr>
<tr>
<td>(b) How much time does it take for the cargo to make contact with the spring?</td>
</tr>
<tr>
<td>(c) What is the work done by air resistance acting on the parachute as it drops?</td>
</tr>
<tr>
<td>(d) What is the average force of air resistance acting on the parachute as it drops?</td>
</tr>
<tr>
<td>(e) None of the above.</td>
</tr>
</tbody>
</table>

6.5.4 Results – Non-traditional problems on examinations

Each exam included a text editing, physics jeopardy, and problem posing task at the end. These problems were assigned for extra credit and presented in a multiple choice format. The rest of the exam contained traditional word problems designed by the instructor. All results from the statistical analysis on the examination data may be found in Appendix E.

On each exam we compared the performance of our cohort group with the rest of the class on each non-traditional problem based on a logistics test using a binomial model. We also compared the performance of our cohort group with the rest of the class on each exam for all of the traditional problems using an ANOVA single factor test with an \( \alpha = 0.10 \) level of significance. The ANOVA was chosen last semester because it is a
reasonably strong test for equality of population means and our data are not violated by the conditions which could threaten the validity of the statistical result. Exam-by-exam and exam-by-treatment interactions were also analyzed, but such analyses required the loss of 42 students that did not take five out of five examinations since both these analyses required mean scores on given problem types for all five examinations. Exam-by-treatment interaction analysis investigated how the treatment and rest of the class varied in performance with respect to the overall averages from all five examinations on problem posing, physics Jeopardy, text editing, and traditional problems. The exam-by-exam interaction analysis investigated how course participants varied in performance between overall examination averages. Exam-by-exam interaction analysis did not differentiate between our treatment group and the rest of the class. Our samples were normally distributed, there was homogeneity of variance, and assumed independence within and between treatment groups. For this semester, our sample was often borderline, not normal. However, the ANOVA is robust with respect to departures from normality. It is only when data is appreciably not normal, that we should worry (Jaisingh, 2000).

The significance level or alpha (α) is the probability of falsely rejecting a true null hypothesis, and is set by the researcher. Because we are comparing teaching treatments, a null hypothesis would mean no difference between the treatments on students’ performance. Thus, if two treatments are not different and the alpha is set to 0.10, then there is a 10% chance of declaring the two treatments to be different when they should be declared the same. An alpha value of α=0.05 is used often based on nothing more than tradition. For this project, an alpha level of α=0.10 was deemed acceptable because our purpose was to show evidence of significance with a small (12 student) treatment sample (McCall, 2001). Often 0.10 alpha level is selected for small pilot projects to determine whether it may be more worthwhile do to a more expansive study. At the α=0.05 level of significance we could make the argument that failing to find statistical evidence that there is a difference does not constitute no difference between groups. Similarly, the α=0.1 level might result in a higher probability of a false positive, but the affect we are looking to achieve from this analysis is not to prove that our focus group intervention was definitively ready for classroom integration, but that there exists some difference between
the groups that could be due to the focus group intervention. Even with a higher significance level, the ‘significance’ doesn’t designate whether the difference is large or important, only that it has a certain probability of existing.

The initial problem-by-problem analysis is described directly below, followed by the exam-by-exam and exam-by-treatment interaction ANOVAs. It is also important to note that using a standard three-factor ANOVA in this case was not appropriate because it failed to model the correlation between the repeated measures. The data violate the ANOVA assumption of independence. Achieving independence would require separate participants taking separate exams, or even separate participants solving separate problems on each exam.

Table 6.11 below shows the comparison of performance on traditional exam problems between our cohort group with the rest of the class on each exam, showing the mean and standard error on each exam. We find that there is statistically significant difference (at the 0.10 level of significance) between our cohort and the rest of the class on their performance on traditional exam problems for exams 2, 3, and 4. These are the three examinations which are taken between the treatment sessions.

Table 6.11 Comparison of scores of cohort and rest of class on traditional exam problems.

<table>
<thead>
<tr>
<th>Exam #</th>
<th>Group Int. Cohort Mean ± S.E. (N)</th>
<th>Rest of the Class Mean ± S.E. (N)</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>69.3% ± 5.46% (N = 12)</td>
<td>61.3% ± 1.25% (N = 229)</td>
<td>0.1553</td>
</tr>
<tr>
<td>2</td>
<td>72.0% ± 5.23% (N = 12)</td>
<td>57.9% ± 1.14% (N = 253)</td>
<td>0.0088</td>
</tr>
<tr>
<td>3</td>
<td>75.5% ± 6.02% (N = 12)</td>
<td>59.5% ± 1.33% (N = 246)</td>
<td>0.0100</td>
</tr>
<tr>
<td>4</td>
<td>87.0% ± 5.69% (N = 12)</td>
<td>73.7% ± 1.26% (N = 246)</td>
<td>0.0238</td>
</tr>
<tr>
<td>5</td>
<td>82.4% ± 5.35% (N = 11)</td>
<td>78.3% ± 1.15% (N = 238)</td>
<td>0.4568</td>
</tr>
</tbody>
</table>

Exams 2, 3 & 4 are \( \leq 0.10 \)
Table 6.12 below shows the comparison of performance on text editing tasks between our cohort group with the rest of the class on each exam, showing the mean for each exam. We find that there is statistically significant difference (at the 0.1 level of significance) between our cohort and the rest of the class on their performance on text editing problems given on examination 2, 4 and 5. It may also be seen that the problem given for examination 3 was not only statistically not different between groups, but the mean score for examination 3 was much lower than other examination scores for text editing.

**Table 6.12 Comparison of scores of cohort and rest of class on text editing tasks.**

<table>
<thead>
<tr>
<th>Exam #</th>
<th>Group Int. Cohort % Correct (N)</th>
<th>Rest of the Class % Correct (N)</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>58.3% (N = 12)</td>
<td>51.1% (N = 229)</td>
<td>0.6237</td>
</tr>
<tr>
<td>2</td>
<td>83.3% (N = 12)</td>
<td>58.9% (N = 253)</td>
<td>0.0738</td>
</tr>
<tr>
<td>3</td>
<td>25.0% (N = 12)</td>
<td>26.0% (N = 246)</td>
<td>0.9373</td>
</tr>
<tr>
<td>4</td>
<td>75.0% (N = 12)</td>
<td>32.5% (N = 246)</td>
<td>0.0033</td>
</tr>
<tr>
<td>5</td>
<td>36.3% (N = 11)</td>
<td>12.2% (N = 238)</td>
<td>0.0461</td>
</tr>
</tbody>
</table>

**Exams 2, 4 & 5 are ≤ 0.10**

Figure 6.26 below shows the problem given for the text editing task on examination 3. The correct answer to this problem is (d), ‘2.0 m/s’ is irrelevant information to solve the problem. The most commonly chosen answer by students was (e), ‘You need all the information given to solve the problem.’ It is possible that students did not recognize this problem as an impulse problem \( (F_{\text{avg}}\Delta t = m\Delta v) \) or did not recognize that the change in velocity was simply the football player’s final velocity minus his initial velocity, but it is not possible to determine the reason without having asked students directly.
Table 6.13 below shows the comparison of performance on physics jeopardy tasks between our cohort group with the rest of the class on each exam, showing the mean for each exam. We find that there is no statistically significant difference (at the 0.1 level of significance) between our cohort and the rest of the class on their performance on physics jeopardy tasks, except on exam 4 when the students in our cohort group performed significantly better than students in the rest of the class (p value = 0.0879). It is interesting to note that there is a wide non-sequential variance of scores within groups between each examination. It is unknown as to the exact cause of this unusual variance, but it can be seen that the variance is not equal between the non-traditional problems, nor does a pattern emerge between problems types. In other words, if a group does poorly on exam 3 with respect to jeopardy problems, and slightly better on exam 4 with respect to jeopardy problems, the same pattern will not necessarily emerge in text editing, problem posing, or traditional problem performances. It was concluded that since no two examinations covered the same material, and the level of difficulty between problems varied along with the topics covered, this variance may simply be an uncontrollable factor.
Table 6.13 Comparison of scores of cohort and rest of class on physics Jeopardy tasks.

<table>
<thead>
<tr>
<th>Exam #</th>
<th>Group Int. Cohort % Correct (N)</th>
<th>Rest of the Class % Correct (N)</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50.0% (N = 12)</td>
<td>49.3% (N = 229)</td>
<td>0.9647</td>
</tr>
<tr>
<td>2</td>
<td>100% (N = 12)</td>
<td>90.1% (N = 253)</td>
<td>0.1185</td>
</tr>
<tr>
<td>3</td>
<td>25.0% (N = 12)</td>
<td>23.9% (N = 246)</td>
<td>0.9361</td>
</tr>
<tr>
<td>4</td>
<td>33.3% (N = 12)</td>
<td>13.4% (N = 246)</td>
<td>0.0879</td>
</tr>
<tr>
<td>5</td>
<td>63.6% (N = 11)</td>
<td>59.2% (N = 238)</td>
<td>0.7704</td>
</tr>
</tbody>
</table>

Only Exam 4 is ≤ 0.10

Table 6.14 below shows the comparison of performance on problem posing tasks between our cohort group with the rest of the class on each exam, showing the mean for each exam. We find that there is statistically significant differences (at the 0.1 level of significance) between our cohort and the rest of the class on their performance on problem posing tasks on exams 3, 4 and 5 (p value = 0.01 on exam 3, 0.001 on exam 4, and 0.038 on exam 5).
Table 6.14 Comparison of scores of cohort and rest of class on problem posing tasks.

<table>
<thead>
<tr>
<th>Exam #</th>
<th>Group Int. Cohort % Correct (N)</th>
<th>Rest of the Class % Correct (N)</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>33.4% (N = 12)</td>
<td>17.9% (N = 229)</td>
<td>0.2138</td>
</tr>
<tr>
<td>2</td>
<td>66.7% (N = 12)</td>
<td>45.1% (N = 253)</td>
<td>0.1409</td>
</tr>
<tr>
<td>3</td>
<td>66.7% (N = 12)</td>
<td>29.7% (N = 246)</td>
<td>0.0101</td>
</tr>
<tr>
<td>4</td>
<td>75.0% (N = 12)</td>
<td>28.0% (N = 246)</td>
<td>0.0010</td>
</tr>
<tr>
<td>5</td>
<td>63.6% (N = 11)</td>
<td>32.3% (N = 238)</td>
<td>0.0384</td>
</tr>
</tbody>
</table>

Exams 3, 4 & 5 are ≤ 0.10

Based on the data above one can see that students in our cohort group performed better than the rest of the class on all three non-traditional tasks and traditional problems.

For the first semester these focus group learning interviews were conducted, there was no statistically significant difference between our cohort and the rest of the class on traditional problems on any of the five examinations. For this semester, our group learning interview participants performed better on three of the five examinations. There are three arguments that might explain for this difference between Phase II and III. The first argument is that the student population for the second semester contained students that had a higher level of problem solving skill than in the first semester. This is possible, but does not explain why our student cohort only did significantly better after the start of our treatment. The second argument is that the ‘traditional’ examination problems created by the instructor are not ‘plug and chug’ style problems, and thus, significantly more challenging to novice problem solvers. This argument is consistent with the notion that traditional plug and chug style problems are amenable to novice problem solving strategies and are not effective assessment tools for gauging students’ conceptual schema improvement. The third argument is that students only performed better when given extra problem solving practice and instruction than the rest of the class.
Interestingly, students in our cohort also perform better on the non-traditional tasks following the start of the treatments. For text editing and problem posing, we also see that students continue to perform significantly better than the rest of the class even after the focus group learning interviews have ended for the semester. Though it is still possible that the extra practice and instruction do in fact play a part in this difference, it is reasonable to assume the treatment method remains promising.

We also compared exam-by-exam and exam-by-treatment interactions using data from only those students who completed all of the exams. These analyses resulted in a loss of about 42 participants per semester out of a total of 253 participants per semester and a loss of one of our 12 students in the treatment group. These results are broad, relating general trends in mean scores, unlike our previous problem-by-problem analysis which investigates differences between instruction using individual problems on individual examinations. As previously stated for the problem-by-problem analysis described above, both the normality and the constant-variance assumptions required for an ANOVA have not been violated.

In order to complete this analysis, each examination was divided into four types of problems: traditional examination problems, problem posing problems, text editing problems, and physics Jeopardy problems. Average scores, or mean scores, were calculated for each problem type for each examination.

The exam-by-treatment interaction analysis determined whether there were differences between our baseline and treatment groups with respect to mean scores and differences in instruction between our baseline and treatment groups. The exam-by-exam interaction analysis determined whether there were differences in all class participants (baseline and treatment students) mean scores between examinations. Tables 6.15 through 6.17 below contain the p-values and description of whether differences are significance with respect to each problem type.

Table 6.15 below displays whether there is a significant difference between the baseline and treatment groups with regards to mean scores on traditional and non-traditional problems. At an alpha level of 0.10, we observed statistical differences between the baseline and treatment group scores on the traditional, problem posing, and text editing tasks. The baseline and treatment groups showed no statistical differences in
the physics Jeopardy scores. That is not to say that the treatment cohort did or did not improve. In order to achieve significance for this analysis, students in the treatment cohort would have to show a large improvement in the last three out of the five examinations such that they overcame the large standard error from their small population. Significant improvement on the first examination would not be expected because it occurred prior to the start of any treatment. Any significant improvement observed on examination two would be unusual given that the examination occurred after two treatment weeks. Significance in this analysis could also be achieved if students overall means were higher on all examinations in one group than the other. This is the case for these participants given that the significant differences occurred in three out the four problem types including the traditional style problems, and the overall higher scores may be observed in Figures 6.27 through 6.30.

Table 6.15 Significance between the baseline and treatment groups on traditional and non-traditional problem scores averaged over all five examinations.

<table>
<thead>
<tr>
<th>Problem Type</th>
<th>p-value</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>0.0001</td>
<td>Significant</td>
</tr>
<tr>
<td>Jeopardy</td>
<td>0.1476</td>
<td>Not Significant</td>
</tr>
<tr>
<td>Text Editing</td>
<td>0.0184</td>
<td>Significant</td>
</tr>
<tr>
<td>Problem Posing</td>
<td>0.0001</td>
<td>Significant</td>
</tr>
</tbody>
</table>

Our analysis also explored problem type interactions between examinations. Table 6.16 displays the p-values (significance of difference) in average problem scores on traditional and non-traditional problems between examinations. At an alpha level of \( \alpha = 0.10 \), there exists a statistically significant difference between problem score averages on the class examinations. This significance suggests that the average scores for different problem types vary significantly between examinations.

Table 6.16 Significance of difference on average traditional and non-traditional problem scores between examinations.

<table>
<thead>
<tr>
<th>Problem Type</th>
<th>p-value</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>0.0001</td>
<td>Significant</td>
</tr>
<tr>
<td>Jeopardy</td>
<td>0.0000</td>
<td>Significant</td>
</tr>
<tr>
<td>Text Editing</td>
<td>0.0001</td>
<td>Significant</td>
</tr>
<tr>
<td>Problem Posing</td>
<td>0.0067</td>
<td>Significant</td>
</tr>
</tbody>
</table>
Using the split plots shown in Figures 6.27 through 6.30 below, we also observe that the differences in average scores between examinations were not always positive. That is, the average scores for a given problem type are statistically different between examinations, but the averages don’t always improve as the semester progresses. The exam-by-exam interaction analysis does not distinguish between the baseline and treatment groups, as it uses the average scores for the total class population, but we may use the univariate split plots below to uncover more information regarding performance differences between the treatment and rest of the class.

A univariate split plot analysis determined that there exists no significant difference at $\alpha=0.10$ between instruction on mean exam scores for traditional or non-traditional problems. The primary difference between this analysis and our previous problem-by-problem ANOVA was that the univariate split plots investigate differences in problem type averages between examinations 1 through 5 while the problem-by-problem analysis investigated differences in instruction by comparing individual students’ problem type performances on individual examinations. Table 6.17 displays the p-values (significance of difference) between the treatment and baseline groups on traditional and non-traditional problem scores between examinations. Our one-factor ANOVA explained earlier in this section provided some promising evidence of significant differences between the baseline and treatment groups for problem types on specific examinations. These trends did not carry forward to fit the overall average from all five examinations.

<table>
<thead>
<tr>
<th>Problem Type</th>
<th>p-value</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>0.1812</td>
<td>Not Significant</td>
</tr>
<tr>
<td>Jeopardy</td>
<td>0.6352</td>
<td>Not Significant</td>
</tr>
<tr>
<td>Text Editing</td>
<td>0.1297</td>
<td>Not Significant</td>
</tr>
<tr>
<td>Problem Posing</td>
<td>0.5704</td>
<td>Not Significant</td>
</tr>
</tbody>
</table>

We can observe treatment by examination interaction by plotting our mean scores across examinations between the baseline and treatment groups. Theoretically, the generated split plot lines should cross or cross eventually if extended when there is any
interaction. Since our data are not error free, these plots are used as a graphical interpretation of our calculations shown in Table 6.17. Below, Figures 6.27 through 6.30 display the split plots for traditional and non-traditional problem scores. We can see that there was no significant crossover, or difference in trends between treatments across repeated measures.

**Figure 6.27** Split plot of the baseline and treatment groups average scores on traditional problems for each examination.
Figure 6.28 Split plot of the baseline and treatment groups average scores on Physics Jeopardy problems for each examination.

![Phase III - Physics Jeopardy Scores](image)

Figure 6.29 Split plot of the baseline and treatment groups average scores on text editing problems for each examination.

![Phase III - Text Editing Scores](image)
Figure 6.30 Split plot of the baseline and treatment groups average scores on problem posing problems for each examination.

Phase III - Problem Posing Scores

As before, an important caveat in interpreting these data should not be overlooked. There is the possibility that the our ANOVA’s revealed false positives, though there would be no way to discern whether this was true without replicating our study with a much larger population of students where a higher limiting significance level would be more suitable.

6.6 Summary

The goal of this study was to examine student participation in activities designed to facilitate students’ development of conceptual schema in problem solving through appropriate use of case-based reasoning. To achieve this goal we conducted a series of six weekly focus group learning interviews with a cohort of 12 students in an algebra-based physics class over the course of a semester. This study was a replication of the previous semester involving eight focus group learning interviews with a cohort of 10 students in an algebra-based physics class. The following research questions are addressed in this study:
How do students perceive the usefulness of an example problem given a new problem possessing the same physical principle but lacking similar objects and/or orientations?

Students were asked to rate the usefulness of examples given during the focus group learning interviews. On average students ratings were moderately high on the likert scale. Students ratings of the usefulness of a problem were lower in week 4 than any other week. It is unknown as to why this week was lower overall, but it is possible that the lower ranking was due to a difference in mathematical complexity between the example problem and problems A and B. Problem C required the students to break up the velocity into x- and y- components, but problems A and B did not. From the previous semesters’ data, it is also likely that mathematics played a key role in whether the problem was considered useful or not useful. It is reasonable to assume this might be the case for week 4 as well. Orientations were changed in various problems over the semester, but the ratings did not drop significantly for changes in axis. Week 6 is a good example of an example problem which contained an angle not seen in problem A or B. For week 6, the usefulness of the problem did not diminish because the angle was present. The difference between this problem and the problem C used in week 4 is that the angle in week 6 does not require any additional mathematics. The angle is there, but does not change the way in which the problem must be solved.

For the individual interviews, students were asked to determine which of six problems would be best and least useful for helping them in a challenging problem. The challenging problem contained both a spring and a pendulum, both moving with simple harmonic motion. Students chose problems as most or least useful solely based upon whether they felt their difficulties were more closely bound to springs or pendulums. For this task, student choices were clearly dependent on the perceived difficulty of one type of simple harmonic motion over the other. It is important to note that there are differences between the equations used and the calculations often asked for between spring and pendulum problems. Spring problems ask for maximum compressions while pendulum problems ask for angles. The differences in mathematical procedure appear to be significant for students.
Given problems that are deep-structure similar and surface feature different, what similarities and differences do students describe as important to the problem solutions?

During the focus group learning interviews, students primarily cite principles and equations as being most similar between problem pairs. Students also cite the motion of objects, additional mathematics, and axis orientations as the primary differences. Students also continue to mention the quantity that is being solved for as a similarity or difference during the focus group learning interviews, but it never ranked high with respect to its importance for finding a solution.

How does students’ emphasis on similarities and differences change given problem pairs with varying deep structure and surface feature similarities / differences?

The students participating in our focus group learning interviews were able to discern the similarities in principle between two problems in a pair that had facial differences and regard such similarities as important to solving the problem. Given a pair of problems that had facial similarities and principle differences, they emphasized the similarity in facial features. Problems with facial and principle differences were rated lowest, and problems with facial and principle similarities were rated highly.

Our results from the usability task during the individual interview also revealed that student selection of examples may be heavily guided by their own comfort level with a given surface feature. A ‘Spring problem’, for example, may be a common example of simple harmonic motion, but are also sometimes designated more difficult than ‘Pendulum problems.’

How does students’ performance on problem solving differ with respect to the rest of the class?

Our results indicate a promising difference in performance on the course examinations between the focus group learning interview participants and the rest of the class. Our cohort seems to do better overall on both traditional and non-traditional exam problems during and after the treatment. It is difficult to determine whether this difference is truly significant with such a small sample, but the statistical result is promising.
How do our results from the previous focus group learning interview study align with this study?

There were no discrepancies between observations made in Phase II and Phase III. Students continue to show that they can and will emphasize deep-structure similarities when given problems that are deep-structure similar and surface feature different. There is evidence that students in our cohort develop the conceptual and problem solving schemata necessary to complete the non-traditional, conceptually oriented tasks given during the individual interviews and examinations.

6.7 Limitations and Future Work

Again, in spite of these promising results it is important to note that the Phase II and Phase III focus group learning interviews were interventions which occurred outside of the classroom with support from two moderators. It would be appropriate that the next step include integration of the focus group learning interview protocols into an algebra-based physics course. Larger samples of data could be collected and compared with previous semesters’ data. Integration and implementation of this intervention in an actual classroom environment would likely affect time management and students’ perception of the tasks. Such an investigation would require significant collaboration with and buy-in from the course instructor.
CHAPTER 7 - Summary and Conclusions

7.1 Summary of this research study

This research project investigated strategies that facilitate students to effectively and efficiently use solved examples to assist with a new problem (case reuse). The rationale behind this study focused on eliciting better construction and reconstruction of problem solving schemata (mental representations of knowledge elements and their associations). Strategies that were used are recognized to reduce cognitive load on working memory due to the processing of procedural and mathematical rules. A thorough review of literature was conducted prior to the start of planning the phases of the project. The pilot and three phases of this project were all built upon the same foundation: valuing worked examples (Ward and Sweller, 1990; Maloney, 1993), active reflection of case comparison (Chi, Feltovich et al., 1981; Graesser, Baggett et al., 1996; Gentner, Loewenstein et al., 2003; Jonassen, 2006), emphasis on deep-structure elements within problem sets (Catrambone and Holyoak, 1989), and assessing students’ development of problem solving schemata using non-traditional problem tasks. The non-traditional problems used were text editing (Low and Over, 1990), problem posing (Mestre, 2002), and physics jeopardy (Van Heuvelen and Maloney, 1999). As this project adapted, researchers continued to use and update the literature review to inform the next viable research phase. We focused our observations on measurement of schema development, and collected information regarding students’ perceptions of implemented strategies.

Pilot Study

In the initial pilot study, we explored students’ perception and understanding of the purpose of two different problem solving strategies. Introductory algebra-based physics students were given an online extra credit problem-solving assignment. They were randomly assigned one of three problem-solving strategies: questioning, structure mapping, and traditional problem solving. Later, eight student volunteers were individually assigned to work problems using one of the strategies in two sessions of semi-structured interviews. The first session investigated students’ general problem
solving approaches a few weeks after they had completed the online extra-credit assignment. The second session investigated students’ perceptions of problem solving strategies and how they related to the extra credit assignments.

**Phase I – Focus Group Learning Interviews using Structure Maps**

The first phase continued our exploration of students’ perception and understanding of a given problem solving strategy, but only one of the original two strategies used in the pilot could be further explored. It was difficult to assess more than one strategy given the time constraints and limited human resources. The structure mapping strategy was chosen between two of our original strategies because it was better received and less misinterpreted by our pilot cohort. Eleven student volunteers enrolled in an algebra-based physics course participated in the semester long study. These participants met in two groups of five and six students a total of nine times during the semester. During these focus group learning interviews, students were asked to solve a set of similar deep-structure problems and discuss the contrast between each of the problems. The selected problems were variations of problems asked in *Physics: Principles with Applications*, Giancoli, 6th Edition. Students were also introduced to structure maps, or visual representations of the associations between quantities for a given broad concept. Students were asked to use the structure maps deemed appropriate while working out the problems handed out by the moderator. They were given time to use the map in their own way and assistance was only provided when students were unable to help one another. Students were also asked to react to the structure maps and discuss elements of the map they found useful. For an assessment of the strategies, students met with a moderator individually twice during the semester. The individual interviews were conducted at the mid- and end-points of the semester and students were asked to perform non-traditional, conceptually oriented tasks. A baseline group, consisting of 10 volunteers with similar grade distribution, was also asked to complete the interviews. Comparison of student performance on the interview yielded no apparent differences between the two groups. Though students ultimately perceived the structure mapping strategy as very useful for problem solving, it was difficult to determine whether the treatment facilitated case reuse or some other aspect of problem solving. Because
students continued to show no evidence of schema development over a semester long treatment, the second and third phase of this project took a slightly different direction. It was decided that the subsequent phase must explicitly facilitate case reuse.

**Phase II – Focus Group Learning Interviews using contrasting cases.**

For the second phase, ten students participated in eight, 75-minute long, focus group learning interview sessions. The topics in each session followed those currently being covered in the algebra-based physics class all participants were enrolled in. Using analogical reasoning arguments for simple comparison and contrasting of cases, a protocol was designed such that worked examples were introduced along side unsolved problems. Step-by-step guides of problem solving included active reflection of principles involved as well as similarities and differences between the worked example and the unsolved problem. Students given different unsolved problems were also asked to compare and contrast their cases, and eventually pose their own problem which incorporated elements from all problems seen during the treatment for that week. To assess the impact of participation in the group learning interviews on students’ problem comparison, the students were also required to participate in two individual interview sessions, one toward the middle and the other toward the end of the semester. During the individual interviews, students were asked to rate the similarities between pairs of problems. The problem pairs were constructed from problems that had facial (i.e. surface) similarities and differences as well as principle (i.e. deep structure) similarities and differences. During these same individual interviews, students were also asked to choose two problems out of a set of six as least and most likely useful for solving a more challenging physics problem. The six problems were unsolved, but students were asked to make their selections based upon how well each speculative problem solution would be most and least useful as examples. Students problem solving performance and conceptual organization of knowledge were also assessed using traditional and non-traditional problems on their five in-class, multiple choice examinations. Our student cohort was statistically compared with the rest of the class.
Phase III – Focus Group Learning Interviews – Replication of Phase II

For the third phase of this study, the focus group learning interviews were replicated using the finalized research protocol from phase II. A group of 12 students were selected from an algebra-based physics class and six focus group learning interviews were performed over the semester. Students were assessed using similar traditional and non-traditional problem solving tasks on five in-class examinations, and an individual interview was conducted at the end of the semester. The individual interview used the same protocol as the second interview from the phase II study.

7.2 Conclusions

In this section we address each of our primary research questions that were formulated at the beginning of this research study.

7.2.1 Research Question #1: What scaffolding -- queues, hints, activities and other external inputs -- cause students to reorganize their knowledge while problem solving?

The strategies used for the second and third phase of this project showed the most promising demonstration of reorganization of knowledge during problem solving tasks. Students for these two phases were treated using a series of reflections aimed to guide students to communicate the physical principles (deep structure characteristics) of contrasting cases. During focus group learning interviews, students were asked to read through an example and ask questions if they had trouble understanding any part of the problem solution. Once students felt comfortable with the solution, each student was handed a worksheet with an unsolved problem and a guided stepwise process aimed at alleviating the cognitive load associated with procedural elements of problem solving. Problems given to students in the groups were not all the same, but they all shared deep-structure (physical principle) similarities and demonstrated different surface feature characteristics (different objects and different axis orientations). Students were asked to directly communicate the principles involved in their unsolved problem, the irrelevant information possibly contained in the problem statement, the similarities and differences between their unsolved problem and the solved example, and the usefulness of the
example in helping them solve their unsolved problem. Once students both individually and collaboratively discussed each of these points, they were then asked to describe the similarities and differences between the unsolved problems with their pairs. They were also asked to explicitly rank their similarities and differences in terms of their usefulness for solving the unsolved problem. Since students were given structurally similar, but not surface feature similar problems, students would often emphasize the deep structure similarities in their own peer-to-peer discussions. Finally these students were also asked to pose their own problem using elements of all the problems seen in the interview.

It was shown through the examination assessments and similarity rating tasks given in the individual interviews that the focus group learning interview treatment likely produced stronger student emphasis on the associations between principles and concepts of given problems. For the third phase, there were also promising performance differences on traditional, textbook style word problems given on the examinations. This is promising evidence for a robust conceptual and problem solving schemata. The results also showed a limitation of these strategies in that students continue to focus heavily on surface feature characteristics when given problems that were surface feature similar. Even if a problem was principle different, if the surface features like an inclined plane or pulley were similar, then students often cue strong association of similarity and usefulness to those features.

7.2.2 Research Questions #2: To what extent can they utilize this scaffolding to reorganize their knowledge while problem solving?

Each of the three primary phases of this project used a new, small (10-12 student) cohort of algebra-based physics students. Students were given a one semester long, week-by-week treatment in one hour and 15 minute intervals. In as few as 7.5 contact hours (including time to eat pizza or sandwiches) outside of the classroom over the entire summer, students showed evidence of better developed problem solving schemata with positive statistically significant differences in both traditional and non-traditional problem solving tasks. Students were also rating problem pairs of a variety of similarities and differences reasonably closer to expert-like ratings after treatment. Students were
moderately comfortable with the given tasks and little training was necessary for a successful phase II and phase III.

Not all phases were as productive. Students participating in the first phase of our project were asked to use structure maps while solving problems. These interviews were conducted over nine weeks with over nine contact hours in the semester, and students continued to show difficulty using the maps while solving problems even after they were modified to better fit students’ suggestions. Students perceived some of the maps as very useful to problem solving, but there were no measurable improvements of students ability to perform non-traditional problem solving tasks given during the individual interview assessments. It was determined after this phase that our strategy needed to be more explicit about eliciting problem comparison.

7.2.3 Research Questions #3: What are the ways in which the expert-like strategy of asking productive questions about a problem can be assimilated in students’ problem solving repertoire?

For our final two phases, students were asked to actively reflect upon the physical principles involved in problems as well as identify problem similarities and differences between similar deep-structure and different surface feature problems. These questions were assimilated into students’ problem solving repertoire by introducing a step-by-step guided process for active reflection before and after solving the problem. The guided process also broke down the fundamental procedural steps involved in solving physics problems as a way of reducing the amount of load on students working memory. The full tasks were performed in approximately an hour and introduced students to three moderately difficult physics problems. These tasks were initially introduced in focus group learning interviews, but may be also possibly introduced in any student pair collaborative setting.

7.2.4 Research Question #4: To what extent is the strategy of facilitating students to ask expert-like questions in an effective way to help students solve problems in physics?
Students participating in our phase II and III treatments were able to show some positive inclination towards expert-like emphasis on principles and concepts of a given problem. This suggests a more expert-like approach to problem comparison during case reuse. Data also suggest that our cohorts were unable to move past the initial problem of focusing on surface features. Rather, when asked to compare problem pairs with like scenarios (i.e., both problems involve rollercoasters), they focused on surface features of problems with respect to the similarity. These students recognized problem similarity at the principle (and equation) level, but chose to designate such structure as less important when scenario characteristics are alike.

7.2.5 Research Question #5: To what extent do students’ attitudes about problem solving change after experiencing the problem solving strategies promoted in this project?

The strategies used for all three phases were well-received by students. In the first phase, students were given structure maps, or visual representations of quantities and their associations. While working through problems, students marked given and asked for information on the maps, and if at all possible, used the map to create a plan for their mathematical solution. This strategy was not as conceptually oriented or schema oriented as would have been preferred by the researchers, but students complied with the strategy if they found it useful and simply ignored it when they did not find it useful. When the maps were determined to be difficult to use rather than useful for solving a given set of problems. Students were not positively disposed towards learning how to use the maps. Instead, students explained why they chose to ignore the maps, gave suggestions for possible changes, and continued to work through problem solutions. This disposition may have been in large part due to a reasonable comfort with ignoring the strategy when it was found to inhibit the process of solving the problems. Moderators rarely asked students to continue to use maps if they were perceived negatively, only asking the cohort to explain why they felt the maps were not practical. By the end of the study, a few students expressed interest in using the maps for their final examination. Since students were allowed to bring ‘crib sheets’ to the exam, it was asked that all prior maps be posted online for quick and
easy access. To our knowledge, at least two out of 11 of our first phase cohort continued to use the maps after the treatment ceased.

For phase II and III, students were asked to use a step-by-step guided worksheet and a fully solved example to assist with an unsolved problem. The first step that asked students to identify the central principles involved in a given unsolved problem initially made students anxious. Students were then told that they would not be graded or assessed on their correctness. From then onward, students completed the step at the start and it never appeared to cause undue anxiety ever again. The participants were fairly comfortable asking questions when they were unsure of their next step. Students sometimes communicated impatience with the step which required students to identify any irrelevant information given in the problem. If the irrelevant information seemed obvious, students would possibly answer with a hint of sarcasm. In the event that problems were too difficult to solve individually, students often turned to one another for assistance, and finally to the moderators. Since students were often provided with a resolution to their discrepancies within that hour, the subtle apprehension dropped away quickly. It was difficult to determine whether students used these strategies after the treatment took place. One would hope that they took away at least some components of the treatment since there were many aspects, like similarity and difference comparisons, which do not require additional preparation by a moderator or instructor and might have been beneficial to the students. More importantly, a difference in students’ emphasis of deep-structure features during problem comparison is evident even after the treatment was completed.

7.3 Implications for further research

This study has identified a framework for facilitating case reuse using deep-structure similarity contrasting cases and active reflection upon the usability of examples. This work was never intended to be turned into a stand alone curriculum. However, it would be appropriate to devise a more refined framework for these case reuse facilitating strategies using data collected from a larger, global sample. This study was limited to outside interventions using small focus group learning interviews. Volunteers were
selected and paid to participate each week. Though data were collected from the class examinations, and extra-credit was assigned to those examinations at the request of project investigators, it would necessary that these contrasting case strategies be implemented and integrated in a full class setting.

In any situation where students are asked to volunteer for educational studies, there exists a subset of a student population that will never participate. An even larger subset of students will never participate given educational studies which require a long-term, full semester commitment. Though our grade distribution modeled the class distribution in variance and borderline normality, we immediately constrained ourselves to a particular population of algebra-based students.

It would also be appropriate for future assessment work to include a wider variance of similarity rating problems, and possibly a wider variance of problem diagrams. In our study, two problems would share the same picture given with the problem statement even when the surface features were similar and not identical. In hindsight, this might have elicited stronger evidence of similarity between the two problems. It would be preferable to either include multiple similar surface feature problems of varying pictures, or to not include pictures at all.

It is also difficult to determine whether our results would be nearly as promising given normal environmental, instructor, and time dependent constraints. Local (continue to use algebra-based physics students at Kansas State University) and global (branch out to different levels of students and different populations of students at different universities) field testing should be the next successive steps of this endeavor.
References


