

Using Twitter and Network Analysis to Assess the Sophistication of Learners' Thinking

In Studio Physics:

An Exploratory Mixed Methods Study

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Deepa Rajiv Deshpande

August 2018

Copyright © 2018 by Deepa Rajiv Deshpande
All rights reserved.

Dedicated

To My Parents

Prabhakar and Snehal

for the unconditional love, unending care and undying support,
which gave me, Milind and Nitin
a head-start in life.

To My Partner in Life

Rajiv

for the countless dreams, dinners, vacations, and moments we cherish
including this Thesis

To My Precious Daughters

Ruchi and Ramya

for filling my life with pure Joy and constant Pride

To My Son

Rohan for making Ruchi happy

And

To all my students of IIT Campus Junior College (1991 to 2007)
whose struggles and successes in learning physics
have truly made this Thesis possible!

ACKNOWLEDGEMENT

This work would not have been possible without the active support and guidance of many scholars, co-investigators, and participants. I would like to acknowledge their contribution here:

Dr. Lisa Yamagata-Lynch, my committee chair, for not only guiding me through the writing of first, the Proposal, and then this Thesis, but also for the wonderful opportunity of being a part of CHATTER'D - a research group initiated by her, dedicated to the use of activity systems analysis and cultural historical activity theory, for examining complex social phenomena. The experiences of investigating Net -neutrality as a team member, and presenting and publishing research papers, in collaboration, have been instrumental in boosting my confidence to use qualitative data collection and analytical techniques for this Dissertation.

Dr. Barbara-Thayer Bacon, my advisor, and committee member, for introducing me to feminist theories, and equity issues in education, and thus being instrumental in directing my attention and interest towards equity issues in physics education. This Dissertation is also a by-product of my research initiatives in this direction, under her guidance, through which I could identify the need for finding ways to collect, analyze, and quantize, large sample qualitative data, to examine issue of equity. Without her support, the teaching and research assistantships which helped me complete the doctoral program, would not have been possible.

Dr. Marianne-Breinig, who trusted in me, and supported me in conducting two studies including this Dissertation, in her course PHYS222. As the Physics expert on my committee, I am also grateful for the many informal discussions with her, about student and TA difficulties, and what may or may not work, in addressing these issues. Her interest in physics education, while a source of inspiration for me, has been the one single factor, which made the implementation of this study possible.

Dr. Patrick Biddix, who introduced me to mixed methods research, relevant literature, and software packages, and guided me through the process of designing this study.

Dr. Mehmet Aydeniz, who agreed to be the science education expert, at a later stage in the Dissertation, and yet, has been very patient in helping me refine the literature to be reviewed, for this study. I am also grateful for his guidance on what I need to consider, from a critique's point of view.

Dr. Michael Waugh, my advisor and former Chair, who not only helped me through the Comprehensive Examination, but has been a huge part of conceptualizing this Dissertation. I am also grateful to him for his references which helped me gain access to and membership of AAPT.

Dr. Barry Golden, science education expert of my Committee, up to the Proposal stage, whose course, 'research trends in science education' was very helpful in developing and refining my research ideas.

And sincere thanks to the Teaching Assistants of Sections 001 and 003 of PHYS 222 Fall 2017, Brennan Hackett, Austin Schmier, Elias Kokkas, and Erich Bermel, who allowed me to implement this study in four studio sessions, and helped me collect all data.

This work was funded in part by the National Science Foundation (DRL-0918409, DRL-0946372, DRL-1247262, DRL-1418288, DUE-0919347, DUE-1225885, EEC-1232656, EEC-1340402, REC-0347000), the MacArthur Foundation, the Spencer Foundation, the Wisconsin Alumni Research Foundation, and the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison. The opinions, findings, and conclusions do not reflect the views of the funding agencies, cooperating institutions, or other individuals. <http://edgaps.org/gaps/acknowledgments/>

I would also like to acknowledge the researchers at University of Wisconsin, whose explicit guidance and support have made my network and quantitative analyses possible.

Dr. William Shaffer, developer of ENA, whose able and prompt response helped me through a ‘software issue’ as I was concluding my analyses using the ENA Web Tool. I am also grateful for his help with understanding the ‘math’ behind the automated calculations of the ENA package.

Brendan Eagan and Melanie Rodgers, who spent many hours on ‘hangout’, to patiently walk me through the process of using the ENA Web Tool. Without their help, I could not have used the package at all.

Dr. Dave Popp, my advisor at Penn State, who trusted in me, and to who I owe the conceptualization and implementation of my Master’s Thesis - an action research study at a public high school in New Jersey, and my very first research venture in Physics Education and Instructional Technology.

ABSTRACT

‘Studio physics’ learning environments are expected to sophisticate learners’ thinking such that a well-organized network of connections is developed, between concepts, their disciplinary representations, and reality. However, most existing assessments are not designed to explicitly capture the connections made, missed, or the misconceptions in learners’ thinking. Four technologies were interfaced in this study, to develop and test an assessment framework which utilized the lens of epistemic network analysis, to enable the capture and quantization of cognitive connections. A mixed methods research design was adopted to collect and analyze qualitative observations of learners’ thinking, as thirty-eight students from two different Sections of a non-Majors physics course, collaborated in groups of three, to complete inquiry tasks of four different ‘experiments’ covering the content of ‘electric circuits’, during one studio lab session. Technologies used were – Twitter, to receive student responses to probing questions sent to them before and after each experiment; Spreadsheet to qualitatively code learners’ responses; ENA Web Tool 4.0, to model the co-occurrences of codes as network graphs and matrices; RStudio to extract quantitative information about the network structure and compute sophistication measures.

‘Epistemological beliefs about the physical sciences’ survey was used to confirm that participants’ beliefs were novice-like, and therefore, context-sensitive. Qualitative analyses revealed that learners’ epistemic constructions of concepts, differentiated by the type and source of knowledge – canonical physics, concrete-intuitive physics, or abstract-intuitive physics. Network analyses revealed that connections between knowledge elements differentiated by the strength and type of connections – canonical, transitional, intuitive 1, or intuitive 2. Quantitative analyses revealed negative sophistication gains for two experiments, zero for one, and positive gains for only one inquiry experiment. Internal consistency reliability of sophistication measures was found to be high, and discriminant validity of the assessment instrument was established with four existing assessments administered before, during, and after the studio session. This study demonstrated that qualitative observations of learners’ thinking could be reliably transformed into quantitative measures of sophistication of thinking, using the QOT2SM framework. Such assessments are significant for evaluating learner differences in thinking, which in turn have implications for ensuring and inquiring equity in practice and research respectively.

TABLE OF CONTENTS

Chapter 1 INTRODUCTION.....	1
Background and Rationale.....	4
Statement of the Problem	11
Purpose of Study.....	12
Research Questions	12
Significance of Study.....	13
Limitations.....	15
De-limitations	15
Assumptions	15
Definition of Terms	15
Conclusions	17
Chapter 2 LITERATURE REVIEW	18
Models of Cognition.....	19
Models of Epistemology.....	22
Epistemic Framing Theory	25
Epistemic Framing Analysis.....	28
Epistemic Network Analysis	31
Conclusions	35
Chapter 3 METHODOLOGY	37
Participants and Background	38
Research Questions	39
Research Design	40
Procedures	47
Data Collection.....	48
Qualitative Analysis Phase 2: Spreadsheet.....	56
Network Analyses Phase 3: ENA Web Tool.....	62
Quantitative Analyses Phase 4: RStudio	67
Mixed Methods Analyses: Data Integration	71
Validity and Reliability	71
Conclusion.....	74
Chapter 4 FINDINGS AND RESULTS	75
Quantitative Results: Quantitative Instruments	76
Qualitative Results: Twitter responses	81
Summary of Qualitative Findings.....	100
Network Analysis Findings: (QUAL+QUAN).....	102
Summary of Network Analysis Results.....	141
Quantitative results.....	144
Summary of Quantitative Findings.....	155
Mixed Methods Analyses: Data Integration	156
Summary of Mixed Methods Findings.....	165
Chapter 5 CONCLUSIONS AND RECOMMENDATIONS.....	166
Discussion of Main Findings	168
Implications for Practice.....	177
Implications for Research.....	184
Concluding Thoughts: Technology and Assessment.....	189

Enhancing Systematic Instruction Design.....	190
Next Steps.....	192
List of References	193
Appendix.....	211
Appendix 1A: Experiment 1 Lab.....	212
Appendix 1B: Experiment 2 Lab.....	213
Appendix 1C: Experiment 3 Lab.....	214
Appendix 1D: Experiment 4 Lab.....	216
Appendix 2A: Study Information.....	217
Appendix 2B: Informed Consent.....	219
Appendix 3A: EBAPS Survey.....	220
Appendix 3B: EBAPS Scoring Scheme	226
Appendix 4A: A3Home work Quiz.....	227
Appendix 4B: Test 1.....	228
Appendix 5: T-tests and Network Graphs Experiment 1.....	233
Appendix 6: T-tests and Network Graphs Experiment 2.....	239
Appendix 7: T-tests and Network Graphs Experiment 3.....	247
Appendix 8: T-tests and Network Graphs Experiment 4.....	256
Appendix 9: R-Script.....	261
Appendix 10: Excel Formulae to Compute S-scores.....	262
Appendix 11: Levels of Sophistication Assignment Criteria	263
Appendix 12: Descriptive Statistics (Measures of Sophistication)	265
Appendix 13: Correlational Analyses Results.....	269
Appendix 14: Cronbach's alpha Results	272
Appendix 15: Inter-Rater Agreement & Code List	274
VITA.....	276

LIST OF TABLES

Table 3.1Pre-Post questions and concepts inquired, by Experiment	54
Table 4.1 Percent Expert Responses on EBAPS Dimensions	77
Table 4.2 Descriptive Statistics for EBAPS, A3Quiz , Lab 4 Quiz, Lab 4, and Test1 scores	78
Table 4.3 Variable Grouping Table	80
Table 4.4 Shapiro-Wilk Test of Normality	80
Table 4.5 Non-parametric Test for Independent Samples	81
Table 4.6 Question Prompts	82
Table 4.7 Experiment 1 Code Categories and Themes	87
Table 4.8 Experiment 2 Code Categories and Themes	91
Table 4.9 Experiment 3 Code Categories and Themes	94
Table 4.10 Experiment 4 Code Categories and Themes	98
Table 4.11 T-test for Differences (PreQ and Post Q Connections).....	104
Table 4.12 T-test for Differences in Epistemic Networks by Section of Course	111
Table 4.13 T-test for Differences in Epistemic Networks between PreQ and Post Q1	114
Table 4.14 T-test for Differences in Epistemic Networks between Section 1 and Section 3	120
Table 4.15 T-test for Differences in Mean Epistemic Networks for PreQ2, PostQ1, and PostQ2	122
Table 4.16 T-test for Differences in Epistemic Networks of students by Section of Course	129
Table 4.17 T-test for Differences in Mean Epistemic Networks of students for PreQ, and Post Q1	133
Table 4.18 T-test for Differences in Mean Epistemic Networks of Section 1 and Section 3	137
Table 4.19 Node(code) and Edge (connection) Summary by Question and Experiment	145
Table 4.20 Types of Connections	146
Table 4.21 Criteria to Assess Level of Sophistication of Thinking	147
Table 4.22 Summary View of S-level Gains by Experiment.....	155
Table 4.23 Correlation Summary by Experiment and Question	156
Table 4.24 Synopsis of Data Analyses and Findings Presented in This Chapter.....	157
Table 4.25 Percent count of theme occurrences in total responses by question and experiment	158
Table 4.26 Summary of t-test Results for Differences in Networks on Five Quantitative Assessments	163
Table 4.27 Summary of Correlations between Sophistication Measures and Quantitative Assessments	164
Table A5. 1 T-test for Significance of Differences in Epistemic Networks by EBAPS Scores - Low, Medium and High (PreQ and Post Q)	233
Table A5. 2 T-test for Significance of Differences in Epistemic Networks by A3HWQuiz Scores - Low, Medium and High (PreQ and Post Q).....	235
Table A5. 3 T-test for Significance of Differences in Epistemic Networks by Lab4Quiz Scores - Low, Medium and High (PreQ and Post Q)	236
Table A5. 4 T-test for Significance of Differences in Epistemic Networks by Lab4 Scores - Low, Medium and High (PreQ and Post Q)	237
Table A5. 5 T-test for Significance of Differences in Epistemic Networks by Test1 Scores - Low, Medium and High (PreQ and Post Q)	238
Table A6. 2 T-test for Significance of Differences in Epistemic Networks of students grouped by A3HWQuiz scores, Low, Medium, and High. [PreQ, Post Q1, and PostQ2]	240
Table A6. 3. T-test for Significance of Differences in Epistemic Networks of students grouped by Lab4Quiz scores, Low, Medium, and High. [PreQ, Post Q1, and PostQ2]	240
Table A6. 4. T-test for Significance of Differences in Epistemic Networks of students grouped by Lab4 scores, Low, Medium, and High. [PreQ, Post Q1, and PostQ2].....	245
Table A7. 1 T-test for Significance of Differences in Epistemic Networks of students by Low, Medium, and High EBAPS scores [PreQ2, Post Q1, and PostQ2]	248
Table A7. 2 T-test for Significance of Differences in Epistemic Networks of students by Low, Medium, and High A3HW Quiz scores [PreQ2, Post Q1, and PostQ2].....	250
Table A7. 3 T-test for Significance of Differences in Epistemic Networks of students by Low, Medium, and High Lab4 Quiz scores [PreQ2, Post Q1, and PostQ2].....	250
Table A7. 4 T-test for Significance of Differences in Epistemic Networks of students by Low, Medium, and High Lab scores [PreQ2, Post Q1, and PostQ2].....	252

Table A7. 5 T-test for Significance of Differences in Epistemic Networks of students by Low, Medium, and High Test1 scores [PreQ2, Post Q1, and PostQ2].....	255
Table A8. 1 T-test for Significance of Differences in Mean Epistemic Networks of students grouped by EBAPS scores, for PreQ, Post Q1, and PostQ2.....	256
Table A8. 2 T-test for Significance of Differences in Mean Epistemic Networks of students grouped by A3HWQuiz (Attempt1) scores, for PreQ, Post Q1, and PostQ2	257
Table A8. 3 T-test for Significance of Differences in Mean Epistemic Networks of students grouped by Lab Quiz scores, for PreQ, Post Q1, and PostQ2.....	258
Table A8. 4 T-test for Significance of Differences in Mean Epistemic Networks of students grouped by Lab scores, for PreQ, Post Q1, and PostQ2.....	259
Table A8. 5 T-test for Significance of Differences in Mean Epistemic Networks of students grouped by Test1 scores, for PreQ, Post Q1, and PostQ2.....	260
Table A11. 1 Connection strength distribution and levels of Sophistication: Experiment 1(PreQ and PostQ)	263
Table A11. 2 Connection strength distribution and levels of Sophistication: Experiment 2(PreQ, PostQ1, and PostQ2)	263
Table A11. 3 Connection strength distribution and levels of Sophistication: Experiment 3(PreQ2, PostQ1, and PostQ2)	263
Table A11. 4 Connection strength distribution and levels of Sophistication: Experiment 4(PreQ, PostQ1, and PostQ2)	264
Table A12. 1 Descriptive Statistics Measures of Sophistication: Experiment 1(PreQ and PostQ).....	265
Table A12. 2 Descriptive Statistics Measures of Sophistication: Experiment 2(PreQ, PostQ1, and PostQ2)	266
Table A12. 3 Descriptive Statistics Measures of Sophistication: Experiment 3(PreQ2, PostQ1, and PostQ2)	267
Table A12. 4 Descriptive Statistics Measures of Sophistication: Experiment 4(PreQ, PostQ1, and PostQ2).....	268
Table A13. 1 Spearman’s rho Correlation Coefficient values between Sophistication Scores and Quantitative Scores	269
Table A13. 2 Spearman’s rho Correlation Coefficient values between Sophistication Scores and Quantitative Scores	269
Table A13. 3 Spearman’s rho Correlation Coefficient values between Sophistication Scores and Quantitative Scores	270
Table A13. 4 4 Spearman’s rho Correlation Coefficient values between Sophistication Scores and Quantitative Scores.....	271
Table A14. 1 Inter-item Correlation and Cronbach’s Alpha between (CCS, I1CS, TCS); (I1CS, K2C); (I2CS, K3C): Experiment 1	272
Table A14. 2 Inter-item Correlation and Cronbach’s Alpha between (CCS, I1CS, TCS); (I1CS, K2C); (I2CS, K3C): Experiment 2	272
Table A14. 3 Inter-item Correlation and Cronbach’s Alpha between (CCS, I1CS, TCS); (I1CS, K2C); (I2CS, K3C): Experiment 3	273
Table A14. 4 Inter-item Correlation and Cronbach’s Alpha between (CCS, I1CS, TCS); (I1CS, K2C); (I2CS, K3C): Experiment 4	273
Table A15. 1 Inter-rater agreement.....	274
Table A15. 2 List of Codes by Experiment	275

LIST OF FIGURES

Figure 1.1. Three corners of an assessment triangle.....	6
Figure 3.1. ENA-based Assessment Framework.....	41
Figure 3.2. Technology-enabled Assessment.....	41
Figure 3.3 . Diagram of Research Design.....	46
Figure 3.4 Screenshot of learners' Twitter responses downloaded using TWEXLIST.....	55
Figure 3.5 Example responses and 21 code columns for PreQ of Experiment 2.....	58
Figure 3.6 Result of consolidation of 21 columns of Figure 3.5.....	58
Figure 3.7 Result of consolidation of columns in Figure 3.6.....	59
Figure 3.8 Resulting theme categories obtained after consolidation and categorization	60
Figure 3.9 Metadata columns (A to AD) containing student scores on different instruments.....	61
Figure 3.10 Raw data column (AE) containing the text of student's Twitter response	61
Figure 3.11 Code columns (AF to AT) containing the code for each code category	61
Figure 3.12 ENA Login Screen	63
Figure 3.13 . The units and conversation fields get populated with the metadata	63
Figure 3.14 Screen with 'Create' button activated	64
Figure 3.15 . Plot screen which shows the ENA sets created.....	65
Figure 3.16 The network space which shows the plot of centroids of all units selected	65
Figure 3.17 . Plot of the mean Epistemic Network of all units selected.....	66
Figure 3.18 Screenshot shows the rDataset for PreQ of Experiment 1	68
Figure 3.19 Screenshot shows the rDataset for PreQ	68
Figure 3.20 . Screenshot of columns A to AD of Output.....	69
Figure 3.21 Screenshot of columns AF to AM of Output.....	69
Figure 3.22 Screenshot of computed score columns of output.....	69
Figure 4.1. Centroids of cognitive networks of students and Mean Networks for PreQ and PostQ.....	104
Figure 4.2. Epistemic Networks of student F23, F32 and F31 for PreQ and PostQ	105
Figure 4.3. Mean and individual student centroids and mean epistemic network for PreQ.....	106
Figure 4.4 . Individual student models with centroids in Quadrant1, 2, 3 and 4	107
Figure 4.5. Mean and individual student centroids and mean network model for PostQ	107
Figure 4.6. Individual student models.....	108
Figure 4.7. PreQ centroids (left upper graph) and Mean networks by Section.....	110
Figure 4.8. Mean networks and centroids.....	115
Figure 4.9. Epistemic networks of students F33, F14, and F8	116
Figure 4.10. Mean epistemic networks of students for PreQ , PostQ1 , and PostQ2	118
Figure 4.11. Subtracted mean networks of Section 1 and Section 3	119
Figure 4.12. Network Centroids of all students for PreQ1, PreQ2, PostQ1, PostQ2	123
Figure 4.13. Network models of student F8, M2, F37, and F5.....	125
Figure 4.14. Mean Network models of all students	128
Figure 4.15. Subtracted Mean Network models of students in Section 1 and Section 3, ...	129
Figure 4.16. Mean Networks of students (left graphs), and Centroid positions	134
Figure 4.17. Individual networks of students.....	136
Figure 4.18. Mean epistemic networks of students from Section1, and Section 3	138
Figure 4.19 . Frequency distribution of levels of sophistication	148
Figure 4.20. Frequency distribution of levels of sophistication	150
Figure 4.21. Frequency distribution of levels of sophistication	151
Figure 4.22. Percent distribution of levels of sophistication	153
Figure 4.23. Percent distribution of themes of sophistication	159
Figure 4.24. Gains in connection weights	160
Figure 4.25. Comparative chart of mean and median sophistication levels 4.....	161
Figure 5.1. QOT2SM Framework of an ENA-based assessment.....	167
Figure 5.2 Studio physics setup	179
Figure 5.3. Percentage of women graduates by Major.	185

Figure A5. 1. Mean networks by EBAPS overall scores - Low (red), Medium (blue), and High (green), for PreQ (Upper 3 graphs), and Post Q (Lower 3 graphs).	234
Figure A5. 2. Mean networks by A3HW Quiz (Attempt1) Scores- Low (red), Medium (blue), and High (green), for PreQ (Upper 3 graphs), and Post Q (Lower 3 graphs).	235
Figure A5. 3. Mean networks by Lab4 Quiz Scores - Low (red), Medium (blue), and High (green), for PreQ (Upper 3 graphs), and Post Q (Lower 3 graphs).....	236
Figure A5. 4. Mean networks by Lab4 Scores - Low (red), Medium (blue), and High (green), for PreQ (Upper 3 graphs), and Post Q (Lower 3 graphs).....	237
Figure A5. 5 . Mean networks by Test1 Scores - Low (red), Medium (blue), and High (green), for PreQ (Upper 3 graphs), and Post Q (Lower 3 graphs).....	238
Figure A6. 1. Mean Network graphs of students grouped by EBAPS scores, Low (red), Medium (blue), High (green), for PreQ (Upper 3 graphs), Post Q1(center 3 graphs), and PostQ2(Lower 3 graphs).....	239
Figure A6. 2. Mean Network graphs of students grouped by A3HWQuiz scores, Low (red), Medium (blue), High (green), for PreQ (Upper 3 graphs), Post Q1(center 3 graphs), and PostQ2(Lower 3 graphs).....	242
Figure A6. 3. Mean Network graphs of students grouped by Lab4Quiz scores, Low (red), Medium (blue), High (green), for PreQ (Upper 3 graphs), Post Q1(center 3 graphs), and PostQ2(Lower 3 graphs).....	243
Figure A7. 1. Mean Network models of students by Low (red), Medium (blue), and High (green) EBAPS scores, for PreQ2 (upper 3graphs), PostQ2(center 3 graphs), and PostQ2 (lower 3 graphs).....	247
Figure A7. 2. Mean Network models of students by Low (red), Medium (blue), and High (green) A3HW Quiz scores, for PreQ2 (upper 3graphs), PostQ2(center 3 graphs), and PostQ2 (lower 3 graphs).	249
Figure A7. 3. Mean Network models of students by Low (red), Medium (blue), and High (green) Lab4 Quiz scores, for PreQ2 (upper 3graphs), PostQ2(center 3 graphs), and PostQ2 (lower 3 graphs).	251
Figure A7. 4. Mean Network models of students by Low (red), Medium (blue), and High (green) Lab scores, for PreQ2 (upper 3graphs), PostQ2(center 3 graphs), and PostQ2 (lower 3 graphs).....	253
Figure A7. 5. Mean Network models of students by Low (red), Medium (blue), and High (green) Test1 scores, for PreQ2 (upper 3graphs), PostQ2(center 3 graphs), and PostQ2 (lower 3 graphs).....	254
Figure A8. 1. Mean epistemic networks of students (upper 3 graphs) and centroid positions of all students (lower 3 graphs) sorted by EBAPS scores, as Low (red), Medium (blue), and High (green), for PreQ (left graphs), PostQ1(center graphs), and PostQ2 (right graphs).....	256
Figure A8. 2. Mean epistemic networks of students (upper 3 graphs) and centroid positions of all students (lower 3 graphs) sorted by A3Quiz scores, as Low (red), Medium (blue), and High (green), for PreQ (left graphs), PostQ1(center graphs), and PostQ2 (right graphs).....	257
Figure A8. 3. Mean epistemic networks of students (upper 3 graphs) and centroid positions of all students (lower 3 graphs) sorted by Lab Quiz scores, as Low (red), Medium (blue), and High (green), for PreQ (left graphs), PostQ1(center graphs), and PostQ2 (right graph).	258
Figure A8. 4. Mean epistemic networks of students (upper 3 graphs) and centroid positions of all students (lower 3 graphs) sorted by Lab scores, as Low (red), Medium (blue), and High (green), for PreQ (left graphs), PostQ1(center graphs), and PostQ2 (right graphs).....	259
Figure A8. 5. Mean epistemic networks of students (upper 3 graphs) and centroid positions of all students (lower 3 graphs) sorted by Test1 scores, as Low (red), Medium (blue), and High (green), for PreQ (left graphs), PostQ1(center graphs), and PostQ2 (right graphs).....	260

LIST OF SYMBOLS AND ABBREVIATIONS

V	Electric Voltage
I	Electric Current
R	Electrical Resistance
C	Capacitance
Q	Electric Charge
A	Area
D	Distance
PER	Physics Education Research
EBAPS	Epistemological Beliefs About Physical Sciences
HW	Home Work
ENA	Epistemic Network Analysis
SE	Sophistication Elements
SN	Sophistication Networks
SN-G	Sophistication Network Graphs
SN-M	Sophistication Network Matrices
SM	Sophistication Measures
QOT	Qualitative Observations of Thinking
ALE	Active Learning Environments
KE	Knowledge Elements
KSEs	Knowledge, Skills, Epistemologies
KSEIV	Knowledge, Skills, Epistemologies, Identity, Values

CHAPTER 1 INTRODUCTION

At the heart of physics education re-search is a shift in emphasis from what are we teaching and how can we deliver it to what are the students learning and how do we make sense of what they do. In order to make this shift, we need to listen to the students and find ways to learn what they are thinking.” (Redish & Steinberg, 1999, p.3)

This Dissertation research explored how technologies can be interfaced effectively, to enable formative assessment of student thinking, in the context of studio physics environments commonly used in undergraduate physics, to facilitate active learning. It is thus, situated squarely at the intersection of Learning Sciences, Physics Education, and Instructional Technology, and embodies the inter-disciplinary spirit of LEEDS (Learning Environments and Education Studies) – my Doctoral program, whose mission is “to equip candidates with the skills and knowledge necessary, to leverage tools and technologies for developing, mediating, and critiquing learning environments” (LEEDS Rationale, n.d.). It is also inspired by my own interest, and informed by my experiences of, learning physics as a graduate student for five years; teaching physics, as a post-graduate teacher, for seventeen years; and utilizing technology to address issues encountered in the learning and teaching of physics, as a graduate student of Education Technology at Penn State, and of Education, at University of Tennessee, for the last seven years.

Learning and doing physics requires complex thinking, to the extent, that it involves thinking about abstract, spatial concepts, which cannot be completely discerned, and/ communicated, without substantial use of advanced mathematical functions, graphical representations, and schematic/ diagrammatic representations (Kozma, 2003). Access to physics, and success in physics, is therefore, contingent on learners’ ability to make concurrent cognitive connections between relevant concepts, their disciplinary representations, and the real world, on demand (McDermott, 2004). The conscious, or unconscious activation of cognitive connections between knowledge elements, in learners’ thinking, is in turn, determined by their epistemologies – their beliefs, and/ expectations about what counts or justifies as the appropriate response in each context (Redish, 2004). For learners of introductory physics courses, activation of a coherent response to inquiry, and/ problem solving tasks, is not a global phenomenon, meaning that productive activation in one context, may not transfer to another context, requiring the activation of same set of conceptual resources, because their epistemologies have been found to be highly context-sensitive (Hoffer, 2004).

Research-based, technology-rich, active learning environments, are therefore, routinely deployed in introductory university physics courses, to enable and strengthen, relevant cognitive connections, or ways of thinking, by having learners engage in structured inquiry and/ problem-solving activities, cultural practices of the community of physics (McDermott, 2001). Various genres of standardized and customized assessment instruments such as concept inventories, beliefs surveys, tests of mathematical and

spatial abilities, have been developed, and are administered, to evaluate the success of the said instructional environments, in producing desired learning gains. Technology-enhanced classroom response systems such as Clickers, have also gained popularity with physics educators, for formative assessment of learners' conceptual understanding, and skills of applying the said knowledge, to perform problem solving tasks. These assessments are designed to measure learning, in terms of disparate outcomes of learners' cognitive processing of assessment tasks, with a tacit expectation, that these outcomes can be used reliably, to predict and/interpret the underlying cognitive process, and therefore, misconceptions in learners' thinking. However, qualitative inquiries of learners' thinking process, have revealed, that to uncover the roots of learners' misconceptions, or the inconsistency of their response, such as is evidenced through traditional assessments cited above, it is necessary to obtain information about learners' framing of the content, and/context, since learners' epistemologies, influence their learning and response (May, and Etkina, 2002; Lising, and Elby, 2005). But existing formative assessments used in physics classrooms, are not equipped to capture and/ interpret such qualitative information required to uncover the underlying cognitive process, and cognitive connections, in learners' thinking.

Physics Education Researchers (PERers) have traditionally relied upon contextual interviews and think-aloud protocols (Hammer, 1994; Lising, & Elby, 2005, Kohl, 2007), video recording of learners' activities (Lising, & Elby, 2005; Scherr, & Hammer, 2009; Bing, 2008; Tuminaro, 2007; Kohl, 2007), journal reflections (May, & Etkina, 2002), and document analysis of students' work such as assignments and quizzes (Fredlund, 2015; Larkin, 2011) to collect qualitative data on learners' thinking. Such qualitative data has been analyzed using the lens of epistemic framing theory, which says, that learners' responses to learning tasks are determined by their epistemic frames – the set of conceptual (knowledge, skills, strategies) and epistemological (beliefs, values, expectations, justifications) resources activated in given learning contexts (Redish, 2004; Tuminaro, & Redish, 2007; Hammer, & Elby, 2003; Shaffer, 2006). However, these data collection and analytical techniques are unsuitable for formative assessment of learners' thinking, in classrooms contexts, for the following reasons: One, the data generated requires cumbersome pre-processing, and digitization, prior to thematic coding and qualitative analyses; and, Two, the models of thinking generated, using epistemic framing theory are essentially qualitative, and will need to be quantized or transformed appropriately, to enable statistical aggregation, description, and computations – necessary features of classroom assessments.

Researchers of the Epistemic Games Research Group at the University of Wisconsin have successfully developed and tested a novel approach called epistemic network analysis (ENA), and designed a digital tool, 'Epistemic Network Analysis (ENA) Web 4.0', which models the associations or

connections between the cognitive elements - knowledge, skills, beliefs, values, and/identities found to co-occur in learners' thinking, as (epistemic) networks, and models learning as shifts or changes in learners' cognitive (epistemic) networks over time. It has been effectively utilized by researchers and educators, in the contexts of virtual internships and games, to assess the sophistication of learners' thinking in-context (Shaffer et al., 2009; Shaffer & Graesser, 2010). Digital artefacts, such as chat room logs, journals, and/reports, have been the primary data collection tools of choice. However, there are no reported instances of its usage in studio lab environments in undergraduate physics, for formative assessment, perhaps for want of tried and tested - data collection mechanisms, and statistical computing tools, which can complement ENA analyses. In physics, only one instance can be found in literature, of the usage of ENA technique to assess learning gains in thinking, in the context of using computational physics to enhance learners' problem- solving skills. The sample size was small, and contextual interviews were used to collect data.

To explore the feasibility of using ENA to monitor shifts in learners' thinking, during inquiry activities, in studio lab settings in physics, there existed a two-pronged need – one, to identify and explore, alternative data collection mechanisms suitable for collecting large sample qualitative data in physical classroom settings; and two, to identify and explore the use of statistical computing package in tandem with ENA, to quantize and compute gains in sophistication of thinking. The primary purpose of this Dissertation research, was therefore, to fill this gap in research, by developing and testing, a technology-enhanced, ENA-based assessment mechanism, which utilized Twitter, to collect qualitative data, ENA Web Tool, to model the coded data as networks, and RStudio, to extract actionable quantitative information from the networked data. In this study, I examined in-depth, how thirty-eight undergraduate students, taking an introductory physics course at a large public university, framed, reasoned and processed cognitively, technology-enhanced, hands-on inquiry tasks, of four experiments covering different, but related content areas in physics, during a three-hour long studio session, by conducting mixed methods analyses of their twitter posts, in response to probing questions, posed at the beginning, and towards the end of each activity.

Qualitative theory-driven content analysis was used to identify learners' epistemic frames, and network analyses were used to model and analyze the resulting epistemic networks graphically, algebraically, and statistically. In the following paragraphs, I will first outline the background and the rationale which ground this study, and then define the research problem, research purpose, research questions, significance of research, its limitations, delimitations, assumptions and key terms. Ideas presented here, outline the theoretical and empirical foundations, which will be developed in further details in Chapter 2. While also supported by literature in Physics Education Research (PER), many of the

ideas presented in the following paragraphs, on common student difficulties, what is involved in learning physics, and what constitute effective practices in learning, teaching and assessment, are based on my own learning experiences of university physics as a graduate student for five years, and on classroom observations compiled through teaching five introductory physics courses including Laboratory Physics, which covered all content areas of domain knowledge – Mechanics, Heat, Sound, Electricity, Magnetism, Optics and Modern Physics, for eighteen years.

Background and Rationale

Discipline based education research (DBER) in the discipline of physics, emerged as a distinct strand of research in the early seventies, in response to an urgent demand for increased participation in physics, spurred by the Sputnik crisis in American History (Docktor, & Mestre, 2014). Thus, initial efforts remained focused on identifying student difficulties, and developing effective instructional strategies to address them (McDermott, & Redish, 1999; Cummings, 2011; Docktor, & Mestre, 2014). Researchers found that traditional methods of instruction wherein concepts were ‘transmitted’ through lecture, and laboratory skills were taught through mere replications of practical procedures prescribed by the instructor, yielded poor conceptual gains (Sharma, Johnston, Johnston, Varvell, Robertson, Hopkins, Cooper, & Thornton, 2010; Hake, 1998). Learners often lacked - the ability to make connections between concepts, their disciplinary representations, and the real world; a coherent and stable conceptual framework; and a developed reasoning ability (McDermott, 2001). Traditional testing based on quantitative problems which required mechanistic application of algorithms was found to be an inadequate criterion to assess functional understanding of physical phenomena (Sharma, et al., 2010).

As evidence began to pile up against traditional methods of instruction, the focus of PER shifted from content/instructor-centered, to student- centered instruction design, informed by the constructivist paradigm, which then dominated the Cognitive and Learning Sciences. Such instructional environments, grounded in the belief, that students demonstrate best learning gains when they are actively involved in the process of knowledge construction, are commonly referred to as active learning environments in physics. The terms active learning (AL) and interactive engagement (IE) have been used interchangeably in PER literature, and are variously described as ‘engagement of learners in activities which require them to predict-observe-explain’, ‘engagement in reflection and expression of reasoning’, ‘interaction with peers’, and ‘intellectually engaged and actively involved in learning’ (Michael, 2006). Effective active learning environments have been known to satisfy two to three criteria listed below (Meltzer & Thornton, 2012, Hake, 1998; Laws, Sokoloff, & Thornton, 1999):

1. Students learn concepts before they come to class and work on inquiry and/ problem solving activities individually or collaboratively during class time,

2. Students are required to think about and articulate their understanding and reasoning,
3. Formative assessments are used frequently during class to get feedback on students' understanding,
4. Use of evidence-based instructional strategies, technology and assessments
5. Use of Peer Instruction, mentoring and collaborative activities

Typical Technology-enhanced Active Learning Environments

Contemporary undergraduate physics courses across the US, commonly employ a combination of one or more research based instructional approaches, e.g. SCALE-UP (Student- Centered Activities for Large Enrollment Undergraduate Programs), Peer Instruction, Interactive Lecture Demonstrations, Tutorial Physics, Workshop Physics, and Real Time Physics, tailored to meet the enrollment size, student needs, syllabus coverage, and pace of the course (Meltzer, and Thornton, 2012). For instance, the 'Elements of Physics' course proposed to be the context of this study, incorporates the IE and constructivist inquiry approaches administered through a combination of two instructional formats: a longer small group (40 to 45 students) inquiry-based studio lab session supervised by teaching assistants; which is preceded or followed by a shorter large class (150 to 200 students) IE lecture session taught by the instructor. Students are expected to read the instructional material and complete homework assignments and online quizzes before class.

During the lab session, mini-lectures are followed by collaborative virtual and hands on inquiry activities designed to include learners with different learning preferences. The activities are structured to have students predict-observe-analyze phenomena under investigation and explain their reasoning. Visualization technologies such as Applets, Simulations and Spreadsheet are integrated to generate dynamic external visual representations of the physical phenomena being investigated, and to vary parameters of inquiry in a way that is not possible in hands on inquiry. Efficient knowledge construction is enabled by supporting learners' abstract visualization process and having students model an expert like reasoning process facilitated by the structured learning sequence (Keiner, & Burns, 2010). Multiple representations of knowledge, are incorporated into activities of both sessions.

Formative assessments, a qualifying characteristic of active-learning environments, are administered using closed-book, and/open-book, in-class, and/ home-work quizzes, with single and/ multiple attempts. Large lectures use interactive-engagement assessments, in the form of class room response systems (CRS), such as timed, clicker questions (small multiple choice conceptual problems), answered in consultation with peers. Responses are auto-scored, and aggregated by the CRS, to generate a graphical output, which is used by the instructor, to lead a classroom discussion and analysis of learners' responses and clarify misconceptions. Labs (reports and/logs of experiments), in studio settings, are also

used as interactive engagement assessments, conducted in groups, and reported in groups/individually. They are scored manually by TAs, with corrective feedback, in the form of written comments, on learners' misdirected conclusions. However, if the feedback could be given in-context, on the lines of Clickers, by administering some form of formative assessment, as proposed and explored in this study, it would be more constructive, in shifting, and/redirectiong learners' framing of content, and activities.

The Assessment Triangle

In a 382-page report titled 'Knowing What Students Know: The Science and Design of Educational Assessment', the committee on Foundations of Assessment, commissioned by the National Research Council (2001), to explore the implications of advances in cognitive sciences, for improving assessments in science and math education, strongly recommended, that technology-enhanced formative assessments which tell us what students know, should be an essential component of student-centered instruction (Chapter 7, NRC Report, 2001), commonly used in science and math. In a student-centered undergraduate physics classroom, many of the learning activities themselves can become valuable instruments of formative assessment, with the use of appropriate techniques of interpretation (Kober, 2015; Singer, Nielsen, & Schweingruber, 2012). The information gleaned from an assessment basically depends on three elements, which form the three corners of an assessment triangle (Figure 1.1 below): (i) the theoretical model of how learners represent knowledge and develop expertise in the subject domain – the model of cognition; (ii) the tasks or situations which allow instructors to observe learner performance – observations; and (iii) the method of interpretation of observed data (Kober, 2015; NRC 2001).

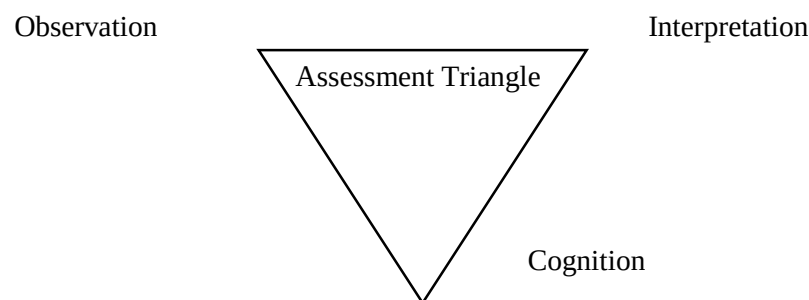


Figure 1.1. Three corners of an assessment triangle.

Issues of Cognition and Interpretation

Effective formative assessments administered during instructional activities, in student-centered learning environments, can reveal students' thinking to the instructor and to students, but their ability to do so, is constrained by the types of tasks or observables available, the model of cognition and the tools chosen to interpret or make sense of the observed data and transform them into measurements or evidence (NRC 2001). For instance, inquiry tasks in studio Lab settings in physics, are commonly designed to seek evidence of learners' understanding and reasoning, through probing questions, which ask learners to describe-explain-predict-justify their findings. While the richness of observations, which consist of learners' Lab/inquiry reports, logs, and/journals, depends on the probing quality of questions, the richness of evidence extracted, depends on the interpretive tool used to assess the observables. Labs are commonly interpreted through the lens of pre-constructed rubrics, but mostly through instructor's subjective interpretation of learners' report.

Interpretation. The rubric used for assessment of Lab reports generally assigns two criteria to assess or interpret learners' competency to inquire and reason - 'completeness' of assigned tasks and 'adequacy of reasoning or explanation' provided by the learner, on three levels of competency - '0 (not attempted)', '75 (inadequate or partial)' and '100' (adequate) in each criterion. Thus, the richness of data contained in the Lab reports is lost in the process of interpretation, because the evidence (evaluation report) does not yield much useful or differentiated information about learners' expertise of reasoning. The instructor's written or oral feedback to the learner, which is based on his/her ideas of cognition and learning, is a richer source of information to the learner about his/her reasoning process, but not amenable to too cumbersome to aggregate qualitatively for the whole class. The nature of evidence gleaned from the data is a function not only of the analytical methods chosen for interpretation, but also, of the theories or perspectives of cognition, which frame the interpretation process.

Cognition. Two distinct theoretical models of cognition have gained prominence in the domain of PER. One is variously described as 'schema theory' or 'mental model theory' or knowledge-as-concepts model proposed by the 'conceptual change' advocates (Bao, and Redish, 2006). The second model is a more recent development, which grew out of a need to 'explain' the process of 'conceptual change' - to identify strategies which could be adopted to replace or change misconceptions or alternative conceptions of learners, constructively, in many learning contexts in physics. The idea of knowledge-in-pieces, or knowledge-as-resources, or phenomenological primitives (p-prims) which are context sensitive and therefore exhibit local coherence rather than the global coherence expected of stable formulations or concepts, was then put forward, and is now categorized as the 'resources model' of cognition (diSessa, 1993; Sherin 2000).

Epistemological component. Alongside conceptual knowledge, last two decades of PER have seen growing evidence of an epistemological component to learning and cognition, which is distinct from the affective component. The affective domain is concerned with feelings, emotions, and factors in the external surroundings which affect learning. While the epistemological domain deals with learners' ideas, beliefs, views, expectations, values, and/interpretations about the nature of physics knowledge, and what justifies as valid sources of physics knowledge, and physics learning - collectively referred to as learners' personal epistemologies or even as epistemic mindsets, which are found to influence their interpretation, reasoning and approach to learning tasks (Ding, 2014; Linn & Songer, 1993; Elby, 1999; Halloun, 1998; Hammer, 1994). Just like the dichotomy of 'modular' and 'model' based accounts of cognitive organization of conceptual knowledge, researchers are also divided in their theoretical accounts of epistemological beliefs – first account being, that they exist as sets of stable, coherent formulations (Hofer, & Pintrich, 1997), analogous to schemas and mental models, and the second account being, that they exist as context dependent epistemological resources (Hammer and Elby, 2003) analogous to p-prims.

Most summative and formative classroom assessments used in physics, whether standardized concept inventories, beliefs surveys, or customized multiple-choice quizzes administered using Clickers or other formats, and/free response tests, are based on the stable, coherent sets of 'concepts and/ beliefs' model of cognition, and therefore are designed to observe singular data about the target of observation – knowledge (and skills) or beliefs. The theoretical lens constrains the observables and/ the interpretations. The 'concepts and/ beliefs' model fails to 'capture' the intuitive and primitive 'pieces' or 'resources' of knowledge and beliefs which could have influenced the unproductive reasoning process reflected by learners' wrong or inadequate responses to assessments. Lab reports discussed earlier, are cases in point of this argument. If the resources model is applied to the same set of data collected by say the concept inventory, the beliefs survey, or the Lab Reports, different, quite often revealing interpretations are generated (Yerdelen-Damar, Elby, & Eryilmaz, 2012).

Epistemic Framing Theory as a Model of Cognition

The Epistemic Framing Theory has emerged recently in PER as a resource-based approach to modeling cognition. Per this theory, learner's thinking, cognitive processing and approach to learning tasks is determined by his/her epistemic frame (Tuminaro, & Redish, 2007; Scherr, & Hammer, 2009; Lising, & Elby, 2005), which comprises of the set of knowledge, skills and beliefs, activated consciously or sub-consciously in learners' memory, in the context of task performance (Redish, 2004; Hammer, 2009). Researchers have used this theoretical lens in physics, to model learners' thinking in various contexts of inquiry and problem-solving activities, using traditional techniques of interviews and

observations to collect data on learners' thinking in-context. The specific elements which constitute learners' epistemic frame - the knowledge, strategies, beliefs, values, identities and/ epistemologies which become the targets of observation, are context specific, and can vary based on the researchers' area of focus. While the data collected, or the observations are qualitative, the methods of interpretation have ranged from purely qualitative thematic analyses, to mixed methods techniques such as epistemic network analysis.

Epistemic Network Analysis (ENA) as a Method of Interpretation

ENA is the network analysis of learners' epistemic frames (Shaffer, Hatfield, Svarovsky, Nash, Nulty, Bagley, Frank, Rupp, & Mislevy, 2009). It can be viewed as the analysis of network representations of cognitive connections enabled (cued or activated) in any given learning context, between the elements of knowledge, skills, beliefs, values, identities, and/ epistemologies that are found to characterize learners' epistemic frames in that context. Thus, the ENA Web Tool, a software package developed by a team of researchers at the University of Wisconsin, models the frame elements of learners' epistemic frames as nodes, and their co-occurrences as connections between nodes, to represent learners' thinking as epistemic networks, and to model learning as shifts in epistemic networks. Per the developers, it provides a 'transformative framework' to assess the development of complex STEM thinking ("Epistemic Network Analysis", 2017).

Digital records of web-based activities such as reading, analyzing and writing research reports; designing and performing experiments; proposing and justifying design prototypes; chat board conversations between student group members; email conversations between student and mentor, were used as data in the context of evaluating the effects of virtual engineering internships such as 'Nephrotex' (Arastoopour, Shaffer, Swiecki, Ruis, & Chesler, 2012), and epistemic games such as 'Digital Zoo' and 'Urban Science' (Beckett, & Shaffer, 2005; Nash, & Shaffer, 2008; Nulty, & Shaffer, 2008; Svarovsky, & Shaffer, 2007), on learners' 'ways of thinking'. A recent application of ENA in the context of an educational game Land Science - an online urban planning simulation, utilized chat conversations (of 155 college students and 156 high school students resulting in 44,964 lines of chat data), along with other work products, and every keystroke and mouse- click, making it possible to assess not only students' final products but also the problem- solving processes they use. (Shaffer, & Luis, 2017)

Developing an ENA-based Assessment

Evidence of ENA's success, in assessing the development of complex thinking, in the context of online epistemic games and virtual internships, specifically designed to achieve this goal in STEM disciplines, suggests that ENA could also be utilized in studio lab settings, in physics, to assess the sophistication of learners' thinking – the development of expert-like cognitive connections, over the

course of learners' engagement in technology-enhanced inquiry activities. The basics steps involved in conducting ENA are – (i) qualitative observations of learners' thinking (ii) qualitative coding of 'observed' data to identify the elements of knowledge, skills, and/beliefs, which frame learners' thinking (iii) coding of co-occurrences of knowledge elements and (iv) transformation of the codes into network graphs and matrices based on the frequency of their co-occurrence in selected time frames, or instances. To implement an efficient ENA-based assessment, which can be used in-context, for formative assessment in studio lab settings, there exists a two-pronged need: (a) addition of another step of computation of sophistication measures, and (b) identification and exploration of suitable technology-interfaces to execute the five steps sequentially, and seamlessly.

Issues to consider in choice of Technology. Social media platforms such as Twitter (Hunco, Elavsky, & Heiberger, 2013) and Facebook, or dedicated web applications such as Today's Meet, and ClassQue (Robbins, 2011) have been used to pose questions and facilitate classroom dialogue, for actively engaging students in reflective thinking and reasoning. *However, there exists a gap in Literature, with respect to the use of ENA to analyze such social media generated digital data collected on students' reasoning and/ blogging their thoughts and assess the development of expert-like ways of thinking.* Choices to execute steps 2 and 3, include Spreadsheet, or Qualitative Data Analysis Software (QDAS) packages such as NVivo, QDAMAX, DiscoverText, and Netlytic; and for step 4, social Network Analysis (SNA) packages, such as Pajek, NetDraw, iGraph, and ENA Web tool 4.0. However, most of the network modeling software packages are not built for conducting advanced statistical analyses expected of classroom assessments. *Also, there exists no precedent in Literature, of the use of other stand-alone SNA package with advanced statistical capabilities, such as UCINET, or the use of programming environments of R and Python, to customize and extend the capabilities of ENA Web Tool, Pajek, NetDraw, or iGraph, to assess complex thinking in the context of large samples.*

Technologies chosen for the study. Keeping in view, the availability of a wide range of technologies, which offer contrasting choices such as stand-alone, customized, licensed applications versus versatile, cloud-based, free or monthly-chargeable application services; and easy-to-use, built-in APIs (application programming interfaces) versus customizable API libraries requiring the use of R, Python or other programming environments, the choice of specific technologies, will necessarily be guided by researcher's purpose, level of technological and programming competency, and the technological resources available to the researcher institutionally, or individually. In this study, the overall purpose of development was to enable formative assessments in inquiry and/problem-solving contexts in introductory physics courses, which cater to a population of students having naïve formulations of concepts and epistemologies. Step (i) of the assessment, was implemented using Twitter and its

applications – TweetGuru, and TWExList, to send question prompts, to students’ Twitter accounts as multi-Direct Message Service (DMS), and to retrieve their responses as direct messages, respectively. Spreadsheet, was used to implement steps (ii) and (iii); ENA Web 4.0 to implement step (iv), and RStudio, to implement step (v). Choice of technologies was based on the criteria that - the technology is freely available; has built-in APIs requiring little programming knowledge; is validated through research and has assured technical support; and offers versatility in terms of its access across multiple platforms, and its capability to allow integrations and customization using R or Python in future (Twitter).

Statement of the Problem

Learners taking undergraduate introductory physics courses tend to bring context sensitive, conceptual and epistemological formulations to the classroom, which demonstrate local rather than global coherence in the context of task performance (Elby, and Hammer, 2001). These resources are determined by learners’ prior life experiences in and out of school, and therefore often differ by demographic factors such as gender and socio-economic status of their families (Magolda, 1992; Schommer, 1992). Technology-enhanced structured inquiry activities are routinely employed in undergraduate physics classrooms, to facilitate for all learners, the development of a stable, coherent network of connections between relevant concepts, their disciplinary representations, physical world, reasoning strategies, and disciplinary procedures - collectively described as conceptual resources, and values, beliefs, and epistemologies – categorized as epistemological resources, through active engagement in the act of inquiry (Kober, 2015; Redish, and Steinberg, 1999).

Therefore, it is important to assess the effects of active learning environments on learners’ thinking, through in-context assessment of the network of connections cued or activated (NRC, 2001; Shaffer, et.al., 2009). But existing assessments, which are based on the concepts and beliefs model of cognition, make quantitative and/qualitative, singular assessments of knowledge, skills and /beliefs, and are not adequate to uncover the connections and/misconnections made, or the connections missed or left ‘hanging’ (NRC, 2001; Cowie, 2012). An ENA-based assessment presents a good fit for assessing the development or sophistication of connections in learners’ thinking (Shaffer, et.al., 2009; Bodin, 2012), during inquiry activities in physics, and technologies exist, to implement it in these settings. However, there exist no examples in PER Literature, of systematic efforts to identify and interface existing digital technologies to develop, achieve and validate, seamless execution of an ENA-based assessment.

Proposed ENA-based assessment. This Dissertation study therefore proposed to explore the reliability and validity of an ENA-based assessment system, represented as Twitter → Spreadsheet → ENA Web tool → RStudio. As the interface representation suggests, the assessment system integrated the technologies of (a) Twitter, TweetGuru, and TWExList to collect qualitative data on student thinking, (b)

the spreadsheet application of MS Office, for qualitative coding of Twitter posts, and for organization of imported and/ exported data (c) the dedicated Web 4.0 ENA tool for generating network graphs and ENA sets, composed of adjacency vectors and data frames, and for conducting network analyses of the coded data, and (d) the open source statistical computing application, RStudio, for extracting quantitative data required to operationalize the construct of sophistication of thinking.

Purpose of Study

The broader purpose of this study was two-fold. One, to expand upon current research base on the use of epistemic network analysis for analyzing complex thinking in STEM disciplines, by examining the performance utility of the proposed ENA-based assessment system, in previously uncharted contexts of inquiry learning in undergraduate physics. Two, to explore how existing, open source and commonly available technologies can be interfaced, to implement an ENA-based assessment, which can maximize instructors' ability to glean useful information about learners' thinking and learner differences in thinking, for making informed decisions about instruction modulation, and inclusive instruction design.

The study had two explicit goals.

- (i) to examine the viability of using Twitter to collect contextual data on learners' thinking, which can generate meaningful information about the composition of learners' epistemic frames after qualitative coding of the data using Spreadsheet,
- (ii) to examine the validity and reliability of the ENA-based assessment, to assess the sophistication of learners' thinking individually, and by aggregate.

The overarching questions which guided the design of this study were:

- How do the inquiry activities compare with respect to their effect on the sophistication of learners' thinking? Do the effects differentiate by gender? (MM)
- Does the proposed ENA-based assessment provide a valid and reliable tool to assess the sophistication of students' thinking in inquiry contexts in undergraduate physics courses? (MM)
- Does Twitter present as a valid and reliable data collection instrument to probe learners' thinking in the context of inquiry activities conducted during studio lab sessions in an undergraduate physics course?

Research Questions

Specific research questions were:

RQ1). What themes of 'epistemic framing' can be discerned from qualitative analyses of learners' Twitter responses, to question prompts sent to them before and after each activity, about the

physical concepts being investigated through each of the four inquiry experiments? How do the pre-and post-activity epistemic frames compare qualitatively? (QUAL)

RQ2). (a) What mean patterns of connections do the Network Graphs and Matrix Data reveal in learners' thinking before and after each of the four inquiry experiments (activities)? (b) Do the pre-and post-activity patterns differ significantly by (i) Section of course and (ii) Levels of (a) Epistemic beliefs Score (b) Lab Quiz Score (c) Lab Score, (d) Homework Quiz Score (Attempt1), and (e)Test1 Scores? (QUAL+QUAN)

RQ3). What levels of sophistication were identified in learners' thinking, through exploratory statistical analyses, of scores of node-centrality and weight density of connections? How do the measures of centrality and connection density of learners' Epistemic Networks relate to (i) Lab Score (ii) Quiz scores, (iii) Test1score, (iv) A3 HW Quiz Score, and (v) EBAPS score? (QUAN)

RQ4). (a) How do the inquiry activities compare with respect to their effect on the sophistication of learners' thinking? (b) Does the proposed ENA-based assessment framework provide a valid and reliable tool to assess the sophistication of students' thinking in inquiry contexts in undergraduate physics courses? (MM)

Significance of Study

This study is significant because of its implications for both research and practice in physics education. Enabling connected ways of thinking and reasoning, development of stable, coherent, expert-like cognitive organization, and beliefs structure, is important for enhancing learners' performance, engagement and participation in physics (Donovan, and Bransford, 2004; Kober, 2015), and it is also the primary goal of design of active learning digital environments commonly employed in contemporary undergraduate physics courses (McDermott, 1990; Meltzer, and Thornton, 2012). But cognitive connections can only be identified using qualitative data collection and analytical methods (NRC, 2001), and qualitative findings must be quantized, to measure the impact of instruction and to enable formative assessment of thinking (Larkin, 2011; Shaffer, et.al., 2009).

A significant contribution of this study was the operationalization of a mixed-methods framework to quantize qualitative data, by integrating existing technologies, to holistically assess, track and monitor not only the conceptual and epistemological resources learners choose to perform specific tasks, but also the structural changes in connections between these resources which are activated while performing inquiry activities. Such a mechanism will enable instructors and researchers to identify specific conceptual, and epistemological issues that maybe preventing learners from making the appropriate connections.

Second, apart from the benefit of providing actionable information for the evaluation and refinement of instruction, the proposed assessment mechanism will also enable instructors to get a deeper and broader view of learner differences in framing, reasoning, and connecting, and thereby to monitor the inclusivity of their instruction. Overall learner participation in physics has remained significantly poor over the years, as compared to other natural sciences (Science and Engineering Indicators, 2014). Besides, females and students from low income families and minority communities remain heavily underrepresented (Hill, Corbett, & St. Rose, 2010; NCES: IPEDS 2011-13).

Most previous works examining the contribution of factors such as learner differences in cognitive skills, styles, beliefs, perceptions, etc., in the propagation and sustenance of inequity of performance and participation, have used quantitative gains in one or more outcomes as benchmarks. These measures are unable to provide the richness of information attained by qualitative data collection and analytical methods, and the qualitative techniques limit generalizability. Mixed-methods mechanisms such as the one explored in this study, will help to bridge the gap between purely quantitative and purely qualitative instruments, and enable the conduct of large sample studies with rich, qualitative data.

Third, the study is significant for research in educational neuroscience, wherein network analyses are the chosen methodology for analyzing neuro-imaging data of cognitive brain function, gathered during performance of tasks in physics. Correlating the network information provided by the EEG and/MRI scans about regions of the brain activated, with the actual thought process, will need simultaneous monitoring of the thought process using technologies which can collect self-reported data in concurrence with the EEG machine, and which can interface with appropriate analytical tools to generate network graphs. Such correlational studies can be operationalized by using the assessment framework tested in this study.

Thus, the assessment mechanism explored through this study is significant for:

- (a) Development of formative classroom assessment tools to assess learners' epistemic network connections in-context, learner differences in epistemic network connections in-context, and therefore to evaluate the effectiveness and inclusiveness of instruction to enhance connections;
- (b) Design of large-sample studies investigating issues of equity which require assessment of learner differences in framing, reasoning and connecting;
- (c) Conducting correlational studies of network graphs of complex thinking process with network graphs of brain scans showing brain activation data in the context of performing tasks in in STEM;
- (d) Validating the resources model of cognition which frames this study.

Limitations

1. Since the study explored new data collection and analytic mechanisms, validation of the assessment instrument could not be achieved fully, and only partial validation was possible, through discriminant validity, but no convergent validity. Limited concurrent validity (elaborated in Chapter 4) was achieved through establishment of discriminatory power of the assessment instrument, in some group comparisons. Follow up studies will therefore, be necessary for validation of the measures and the instrument.
2. Since the study was exploratory, technological and software developments were also exploratory, and further refinements of the Application Programming Interface, and customizations of rScript will be necessary prior to its implementation in other contexts.
3. Since the assessment system, is designed to empower instructors with tools which can be customized per their learning goals, and priorities, replicability can only be possible for the system (or framework), and not specific instance of assessment, which are essentially customizations of the system.

De-limitations

The major de-limitation of this study was that the participants being examined consisted of an intact sample of two sections of the course. Thus, the sample may not constitute an accurate and true representation of the target population of all non-Majors' undergraduate students taking this course. Secondly, the qualitative coding process was informed by researcher's expertise in physics, and long-standing experience as a physics educator, and although efforts were made to conform to the goals identified by the course instructor, the code categories may not truly reflect the instructor's expectations of sophistication.

Assumptions

This study was conducted under two major assumptions. One, that co-occurrence of cognitive elements corresponded to cognitive connections between the elements identified by the qualitative analysis of learners' Twitter reflections. Second, that the information on learners' thought processes as provided by the participants through their Twitter posts, reflects their thinking accurately.

Definition of Terms

The following definitions of terms were adopted for the purposes of this study. They have been synthesized through a review of literature presented in the following section.

- 1) Epistemic frames – The set of cognitive elements cued into activation, consciously or unconsciously, when learners are presented with a task, and which determine how learners interpret, reason and respond to given learning tasks (Bodin, 2012; Redish, 2004; Schaeffer, et.al., 2009).

- 2) Conceptual Resources – Elements of Knowledge, skills and procedures activated when a learner is presented with a learning task. (Redish, 2004; Tuminaro, & Redish, 2007; Hammer, & Elby, 2003).
- 3) Epistemological Resources – Elements of epistemological beliefs, values, expectations about physics knowledge, its valid sources, and about physics learning, which influence the set of conceptual resources chosen by learners consciously, or unconsciously, when presented with a learning task. (Redish, 2004; Tuminaro, & Redish, 2007; Hammer, & Elby, 2003).
- 4) Epistemic framing – The process of interpreting a given learning situation using a set of conceptual and epistemological resources activated, when learners are presented with a learning task. (Bodin, 2012; Redish, 2004)
- 5) Epistemic Network – The virtual model of connections (between the conceptual and epistemological resources) activated in learner's memory, when presented with a learning task. (Bodin, 2012; Schaeffer, et.al., 2009).
- 6) Nodes – The set of epistemic elements which co-occur in learners' thinking and form the nodes of epistemic networks.
- 7) Edges – The set of connections assumed to be made when two epistemic elements co-occur in a time segment of inquiry activity, or are semantically adjacent in the data, and form the edges of epistemic networks.
- 8) Sophistication of thinking – A construct which characterizes the types and strength of connections in learners' thinking, modeled as epistemic network.
- 9) Measures of sophistication – Scores computed from the raw, and/ transformed data of strength or weight of network connections.
- 10) Measures of node centrality (Node Weight) – The weight density of connections to and from any given node in an epistemic network.
- 11) Measures of connection strength (Edge Weight) – The weight density of connections of all, or specific types of connections in an epistemic network.
- 12) ENA-based Assessment – An assessment which uses epistemic network analysis for interpretation of the assessment tasks.
- 13) Inquiry activities – Specific hands-on and/virtual experiments/explorations (investigations) that the students are assigned to complete during each studio lab session.
- 14) Inquiry Lab Reports – Reports of inquiry activities completed by students, which consist of their responses to specific predict-explore-observe-describe-explain-justify questions that are integrated into the inquiry activities.
- 15) Twitter posts – Learners' twitter responses to probing questions administered before, and after each inquiry activity.

16) Studio Lab session – An interactive laboratory session during which learners complete each of the inquiry activities, or experiments, assigned to them, in group settings.

Conclusions

This chapter began with a discussion of the structure of active learning environments commonly deployed in introductory university physics courses to enhance the sophistication of learners' thinking, and the role, and issues of formative assessments used in these settings. The concept of Epistemic Network Analysis was introduced as an alternative to existing assessments, with a brief discussion of the theoretical model of cognition which frames these assessments. A statement of the problem being addressed by the study was presented, along with purpose of the study, research questions, and significance of the study. The chapter concluded with acknowledgement of limitations, delimitations, assumptions, and definition of key terms used in the study. In the following sub-section, an overview of the organization of this Dissertation Thesis will be presented.

Organization of Dissertation Thesis

This Dissertation will be structured into five chapters. Chapter One presented an introduction to the research study, giving a brief background, and stated the problem, and purpose of study. Chapter 2 will review literature on Models of Cognition, Epistemology, and Epistemic Framing theory. Previous works in physics on Epistemic Framing Analysis, and on Epistemic Network Analysis, will be discussed to build up a rationale for the study. Discussion will revolve around 'three corners of the assessment triangle', model of cognition, model of interpretation, and the observables which will be interpreted to extract evidence of learning – sophistication of thinking, in the context of this study. Empirical foundations for the use of Twitter, will also be elaborated.

Chapter Three will present the methodological framework adopted in the design of this study, and the rationale for adopting an integrated Mixed Methods Design. Thus, the research design, and Methods, including Data Collection, Analysis, and issues of Validity and Reliability, for Qualitative, Quantitative and Mixed Methods research will be elaborated in Chapter 3. Chapter Four will present the findings, results, and conclusions of the study, in five sections – quantitative grouping, qualitative results, to answer RQ1; findings of network analyses to answer RQ2; quantitative findings to answer RQ3; and results of mixed methods analyses to answer RQ 4. Chapter Five will present a discussion of future steps and implications for teaching and research.

CHAPTER 2 LITERATURE REVIEW

As elaborated in the previous chapter, overall purpose of the study, was to fill a gap in research, and in practice, by developing and validating a technology-enhanced assessment framework, to assess the sophistication of learners' thinking, in the context of studio physics - research-based learning environments, commonly deployed in introductory university physics courses, to promote active learning. Some of the key points discussed in the previous chapter, about learning and assessment in the context of studio physics, which frame the review of literature presented in this chapter, were as follows: (a) Studio physics is a research-based learning environment, commonly deployed in introductory university physics courses, to promote active learning (Cummings, et.al., 1999). (b) Formative assessments are a defining characteristic of studio physics (Cummings, et al., 1999; NRC 2001). (c) 'Observations' of learners' behavior, performance, and/response; the model of 'cognition' which underlies the interpretive technique; and the model of 'interpretation' used to extract evidence of learning from the 'observations', form the three corners of an assessment triangle, and must be specified or defined while designing an assessment (NRC 2001). (d) Epistemic framing theory and Epistemic Network Analysis present a good fit as model of cognition, and model of interpretation, respectively, for assessing sophistication of learners' thinking (Shaffer, et.al., 2009) in the 'context' of studio physics inquiries.

In this chapter, I will build up a rationale for choosing epistemic framing theory, and epistemic network analysis, as models of cognition and interpretation, respectively, for the design of proposed ENA-based assessment framework, to assess the sophistication of learners' thinking, in studio physics environments. I will first present an overview of assessment of expertise and sophistication in physics, and then present an overview of the models of cognition and epistemology, in two separate sections. In the following section, I will discuss how epistemic framing theory integrates cognition and epistemology, to present a unified theoretical framework of cognition, which was chosen for the design of assessment in this study. Then the models of interpretation – epistemic framing analysis, and epistemic network analysis, will be elaborated through empirical works, in the last two sections, to explain the rationale for choosing Epistemic Network Analysis, as the model of interpretation, for developing the assessment in this study.

Expertise and Sophistication in Physics

What does 'expertise in physics' mean? Traditionally, expertise in physics has been characterized either by the depth, breadth, and organization of canonical knowledge, or by the strategies chosen to carry out problem solving tasks (Redish, 2005). In recent years, Physics Education Research (PER) has seen an increase in focus, on a third criterion of expertise - epistemological sophistication, or learners' justifications and/reasoning, for choosing specific knowledge elements, knowledge representations, and/

strategies to carry out domain-specific tasks (Elby, and Hammer, 2001; Redish, 2005). Conceptual questions, and/problem solving tasks, are used to assess expertise of knowledge, and skills, while surveys of beliefs, expectations, and/reasoning questions are used to assess epistemologies (Dowd, 2012). But all three lines of assessment, essentially strive to measure or examine the sophistication of learners' thinking - the alignment of learners' cognitive processes, such as reasoning, knowledge construction, organization, and/value-estimation, with the epistemic practices (ways of knowing) which are considered norms, or standards adopted by the community of physics (Smith, and Tanner, 2010; Dowd, 2012).

Thus, sophistication of thinking means, achieving close alignment between the ways in which learners connect knowledge, skills and epistemologies (KSEs) consciously, or subconsciously, and the ways in which experts in physics, make these connections to construct a productive response, when presented with problem solving, and/inquiry tasks (Elby, and Hammer, 2001). Research suggests that for novice learners, the activation of relevant knowledge and skills present in memory, is sometimes hindered/constrained by their naïve epistemologies which are cued by the context (May, and Etkina, 2002). Thus, the occurrence of each KSE is an inter-dependent event, and therefore, it is the co-occurrence of KSEs, which should be observed, to assess sophistication of thinking, rather than mere occurrence, as is done by existing assessments commonly used in university physics courses (Shaffer, et. Al., 2009). In the following section, I will elaborate the 'conceptual' component of cognition – the large grain-size account of concept formation as stable structures called schemas, or mental models; the small grain-size account of concept formation as unstable formulations called primitives, knowledge-in-pieces, or intuitive knowledge; and Vygotsky's account of concept formation, as two distinct systems, spontaneous and scientific, which develop in parallel, in opposite directions, and mediate one another.

Models of Cognition

The terms 'Learning' and 'cognition' are used together quite often in the field of education, since the two processes overlap. To explain how students, learn physics for instance, researchers either describe the structure and organization of cognitive elements (knowledge, epistemology, etc.) (diSessa, 1998), and/ identify the cognitive process of organization or construction of this cognitive structure (Bing, 2008, Tuminaro, 2004). Instruction and assessment, which mediate learning, are then developed around the specific models of cognition, adopted (NRC 2001). Most learning environments and assessment instruments used in contemporary classrooms in science and math, are designed on a common belief that learners construct new knowledge from prior knowledge (Bransford, Brown, & Cocking, 1999). However, the cognitive process of knowledge construction, implicated in the process of learning, has been described in myriad distinct ways. Prominent among these, are the theoretical models of cognitive constructivism, and social constructivism, attributed to the fields of Learning Sciences and Education

Psychology (Greeno, Collins, and Resnick, 1996; Schunk, 2012) and the model-based and resources-based accounts of cognition, originating in the realm of science education, and Cognitive science (Redish, 2004). While the cognitive and social constructivist models theorize or explain the learning process in terms of construction of knowledge as being an individual or a social process, the model-based and modular accounts of cognition, focus on the cognitive process of knowledge construction of individual learner, in terms of the structure and/ process of concept formation, including the effect of social, cultural experiences. The model-based and resources-based (modular) accounts of cognition, will be elaborated in the following paragraphs, along with common threads from Vygotsky's theory of concept formation. The resources-based account is a precursor of epistemic framing theory, adopted as a model of cognition in this study, and Vygotsky's theory of concept formation informed the qualitative analyses of learners' thinking.

Two models of cognition, which are distinguished by the size of cognitive structure or knowledge structure used as the unit of analysis –fine-grain or large- grain size (Redish, 2004; Hammer 2000; Taber, 2008), have gained prominence in PER. The large grain size model is termed as unitary or model-based account of cognition and is attributed to the 'conceptual change' and 'misconceptions' theorists, who hold the view that learner's knowledge structures exist as coherent stable concepts (large grain size) which may or may not resemble concept formulation of experts (Carey, 1999; 2000). Per this model of cognition, if learner's prior knowledge formulations differ from those of experts, then conceptual change strategies must be implemented to resolve these differences (Caramazza, McCloskey, & Green, 1981; Vosniadou, & Brewer, 1992). Some other theories which are based on this model of cognition, are schema theory, and the mental model theory, named after the cognitive structure which forms the basic unit of analysis (McCloskey, 1983a, 1983b). The common thread, is a belief in stable cognitive structures, which must be 'confronted' or 'restructured', to change their formulation (McCloskey, 1983a, 1983b).

diSessa challenged the idea of 'concept' and 'conceptual change', and introduced 'coordination classes' –knowledge structures used by learners to 'coordinate' information, and implement 'readout strategies' to 'read', 'see', or 'connect' information, using 'causal nets' – set of inferences about the relevance of information (DiSessa, 1993; Sherin 2000; McDermott, 1990). The coordination class theory has been extensively used in science education research, to characterize learners' thinking process. Misconceptions of novice learners, were described by diSessa, as 'knowledge-in-pieces' - weak, unstable, knowledge formulations, or primitives, or intuitive knowledge. While 'invariance' from one context to another is a defining characteristic of 'coordination classes', context-sensitivity, or local, rather than global coherence, defines knowledge-in-pieces (kip) (diSessa, and Wagner, 2005). Different

categories of primitives have been identified, based on the specific cognitive process implicated in their usage e.g. reasoning primitives, phenomenological primitives, and epistemological primitives. These unstable, loose, abstract formulations of raw intuitions or primitive ideas, are believed to be rooted in learner's everyday interactions and experiences at home, school, class or community (Vosnaidou, 2002; diSessa, and Sherin, 1998).

The small-grain size account of cognition, and its associated cognitive structures – kip, intuitive, or spontaneous knowledge, explains the context-sensitivity of learners' cognitive process, and has been referred to, recently in PER as the modular, manifold account, or the resource-based model (Tuminaro, 2004; Redish, 2004) of cognition. The study of expert-novice differences has revealed that experts in physics demonstrate a coherent, extensive conceptual framework compiled with a range of resources - concepts, schemas, mental models, procedures, reasoning strategies, and epistemologies (Redish, 2004). Schemas (simple mental models), mental models, concepts are robust patterns of association of various conceptual resources, tightly connected, or strengthened with repeated activations, and which activate reliably in a variety of situations (D'Andrade, 1995; Carey, 1999; 2000). In contrast, novice learners demonstrate a weakly organized conceptual framework, made up of fragmented pieces of knowledge elements and resources (diSessa, Gilespie, and Esterley, 2004).

Although the 'coordination class and kip' model offers a way to explain 'conceptual change' better than the traditional 'concepts' model, by formalizing the existence of intuitive knowledge, it does not explicitly offer a mechanism, to reconcile the simultaneous existence of both knowledge structures in learners' cognitive system. Vygotsky's theory of concept formation, and mediation, is more useful in this respect. Vygotsky identifies three phases of development of a child's concept to an adult concept, with each phase, comprising of distinct stages : (i) The concepts in first phase, are called Conglomerates, which are 'vague, syncretic conglomeration of individual objects which somehow coalesce as a highly unstable mental image in a child's mind; (ii) The concepts in second phase are called Complexes, in which objects are not only grouped or heaped together in a syncretic manner, but are also conglomerated on the basis of factual, concrete, connections, bonds, associations between the objects; and (iii) The concepts in the third phase are distinguished by their generalization and abstraction. (Vygotsky, and Kozulin, 1986; Otero, 2004).

Vygotsky further differentiates the development of concepts into two distinct conceptual systems: (a) Spontaneous concepts which are grounded in concrete, informal experiences, and grow upwards to generalization or abstraction of meaning of individual elements which form the concept, and (b) Scientific concepts which develop downward from abstract knowledge articulated using signs and tools shared by the community (graphical, mathematical, symbolic representations), to its concretization, or application to

real world experiences, through formal schooling or instruction (Otero, 2004; Vygotsky, and Kozulin, 1986; Rieber, 1999). According to Vygotsky, although spontaneous and scientific concepts have different origins, they develop simultaneously, mediating the development of one another (Vygotsky, and Kozulin, 1986; Otero, 2004). In the development of Spontaneous concepts, experience-based meaning, precedes organization, articulation in terms of shared language and abstraction of the concept; while for Scientific concepts, articulation and abstraction of the concept precede development of its meaning (Otero, 2004; Rieber, 1999; Vygotsky, and Kozulin, 1986; Derry, 2013). Thus, students may be able to articulate canonical ideas, without complete understanding, and therefore, are unable to apply them to real-world problems, consistently.

While structure, or organization, and origin, of knowledge structures, as discussed in this section, have been used traditionally, to assess sophistication of thinking, a need arose for considering another criterion – epistemological sophistication, as evidence began to accumulate in PER literature, that learners’ expectations and epistemological beliefs influenced their cognitive processing of tasks, significantly (Carey & Smith, 1993; Elby, 1999; Halloun, 1998; Hammer, 1994; Hewson, 1985; Linn & Songer, 1993; Redish, Steinberg, & Saul, 1998; Roth & Roychoudhury, 1994; Smith, Maclin, Houghton, & Hennessey, 2000). Assessment of epistemological sophistication developed as a separate line of research, and different models of epistemologies emerged alongside traditional models of cognition. In the following section, I will review some of the prominent models of epistemology, before elaborating in the following section, the epistemic framing theory, which essentially integrates the resources -based accounts of cognition and epistemology, under a unified framework.

Models of Epistemology

Epistemology from a philosophical perspective, includes theories about what counts as knowledge, and deals with questions surrounding the nature, source, methods, limits, and justification of people’s beliefs and knowledge claims (Hofer & Pintrich, 1997; Thayer-Bacon, 2003). These theories are applied to determine if a belief is ‘mere belief’ or ‘rational belief’ supported by compelling reason, or ‘justified true belief’, which is a necessary condition of knowledge (Thayer-Bacon, 2003). On the other hand, the terms ‘epistemologies’, ‘epistemic beliefs’, ‘epistemic cognition’ (Chin, Buckland, & Samarapungavan, 2011), ‘epistemic mindsets’ as used in field of education, encompass all kinds of explicit or tacit cognitions, ideas, views, stances, positions, related to matters concerning the process of knowledge construction; and embody a psychological approach to the traditionally philosophical field of epistemology. While there is no consensus on the construct itself, or its dimensions and/ factors, through its linkages to other constructs in cognition and motivation, nature of its domain specificity or generality, and its gender effects (Hofer, & Pintrich, 1997), contemporary education research in this area does reveal

three prominent models of (personal) epistemology: developmental, multidimensional beliefs based, and resource-based. These models are elaborated in the following sub-sections.

Developmental Model

Researchers using the developmental model have mainly examined how individuals interpret their educational experiences (Perry, 1970, 1981; Belenky, Clinchy, Goldberger, & Tarule 1986; & Baxter Magolda, 1987, 1992), and how their epistemological assumptions influence thinking and reasoning processes such as reflective judgement (King & Kitchener, 1994; Kitchener & King, 1981) and skills of argumentation (D. Kuhn, 1991, 1993). Their findings suggest a progression of epistemological development from low to high levels of sophistication, and that age, gender and educational experience, influence the pace and stage of development of epistemological position or stance of an individual. The works of feminist scholars, which show that females demonstrate ‘connected ways of knowing’, while males demonstrate ‘separated ways of knowing’, are generally attributed to this line of research, although many feminist scholars do not agree with this consensus view (Belenky, Clinchy, Goldberger, & Tarule, 1997; Baxter Magolda, 1992).

Beliefs Model

Epistemologies as sets of beliefs, or the dimension-based models originated out of a need to segregate the construct of epistemology into multiple dimensions, because researchers (Schommer, 1990, Hofer and Pintrich, 1997) believed that these dimensions evolve and/develop separately, although a relation between dimensions was not ruled out (Duell, and Schommer-Aikins, 2001). Most dimensional models of epistemic beliefs are based on two dimensions in the category of ‘nature of knowledge’: (I) degree to which it is simple, consisting of unconnected facts, or complex, consisting of related concepts, and (II) the degree to which it is certain (unchanging), or tentative (evolving). The other two dimensions relate to ‘nature of knowing’: (III) source of knowledge, whether it resides in authority outside oneself, teacher, text etc. or it is self-constructed, and (IV) justification for knowing or the method of evaluating sources, e.g. whether learners count beliefs and opinions as knowledge, or seek reasoned justification for beliefs (Hofer, & Pintrich, 1997). This model has also been used extensively by researchers to show that learner differences in views/beliefs about the nature of knowledge, learning and what counts as evidence of knowledge and learning are associated with differences in ways of approaching learning tasks, reasoning, as well as interpreting tasks and instructional contexts (Ding, 2014; Hammer, & Elby, 2003), and therefore differences in learning performance (Carey & Smith, 1993; Elby, 1999; Halloun, 1998; Hammer, 1994; Hewson, 1985; Linn & Songer, 1993; Redish, Steinberg, & Saul, 1998; Roth & Roychoudhury, 1994; Smith, Maclin, Houghton, & Hennessey, 2000).

Domain-specificity. Philosophically, criteria for evaluating science claims may be quite different from criteria for evaluating claims in literature. Psychologically, different cognitive processes are involved in learning different domains. Besides, expert knowledge is domain specific, suggesting that different epistemologies apply to different domains of knowledge (Muis, Bendixen, & Haerle, 2006). This has led to the advocacy and development of self-report survey questionnaires which are domain-specific, and the statements are phrased in discipline specific language. Comparative studies across domains show that learners demonstrate different patterns and levels of sophistication in different disciplines, for equivalent dimensions (Buehl, Alexander, & Murphy 2002; Greene, Torney-Purta, & Azevedo, 2008). Most students were found to believe that mathematics knowledge consisted of numbers and operations; while beliefs about the nature of social studies knowledge were found to be diverse, and included notions about people, places, events, and historical periods (Schoenfeld, 1989; Stodolsky, Salk, & Glaessner, 1991; Hofer, 2000; Buehl, Alexander, & Murphy, 2002). Hence domain-generalism is no longer considered a viable option, although there are differences on the degree of domain-specificity, and some researchers have suggested levels of epistemic cognition, from domain-specific, to academic and then general (Hofer, 2006; Muis, Bendixen, & Haerle, 2006).

Domain-specific epistemic beliefs are shaped by the instructional environment, which includes grading and practices (Schoenfeld, 1989; Stodolsky et al., 1991). Domain specialization typically begins in upper level high school, and students' domains of study through college and graduation, determine their epistemic beliefs (Muis, et al., 2006; Royce, 1978; Hofer 2000). However, the correspondence between students' domain specific epistemic beliefs and those of experts does not become apparent until they complete graduate school (Muis, et al., 2006). Domain specific beliefs at the undergraduate level are largely context-dependent, and naïve; thereby context in one domain may stimulate cognitive resources relevant to a similar context from another domain, meaning that epistemic beliefs of one domain may influence (positively or negatively) learning in another domain (Muis et al., 2006; Hofer, 2000; Schommer, Dual, & Barker 2003). Such cross referencing is likely to happen more often at the undergraduate level, when students are undecided about their majors; and a negative interaction between learner's epistemic beliefs in one domain with another domain's instructional environment, may dissuade students from persisting in further studies in the latter domain (Muis et al., 2006; Hofer, 2000).

Resources Model

Hammer and Elby (2003) are the main proponents of the resources model of epistemologies, which was developed to resolve the conflict between learners' epistemologies as measured by using the beliefs survey, and epistemological views demonstrated by learners through in-context interviews. Learners also tend to adopt different epistemological stances, based on context, and these different stances

work productively for them in learning concepts in physics (Elby, and Hammer, 2001). For instance, learners' belief that knowledge is tentative and evolving, may not be useful to the learner, when applying Newton's third law to solve problems, but it will be useful in understanding motion in non-inertial frames of reference (Hammer, 2000). Some of these productive stances can be used as 'epistemological anchors', to redirect learners' conceptual formulations (Hammer, and Elby, 2003; Hammer, 2000). Thus 'epistemological resources' are the resource equivalent of the unitary concept of epistemological beliefs.

The beliefs and resources models of epistemology align in parallel, with the concepts and resources accounts of cognition. Thus, the model-based, or concepts and beliefs account of cognition (knowledge exists as models, concepts, and epistemological beliefs) provides a useful analytical tool, when learners show evidence of consistent coherence of conceptual and epistemic formulations in different contexts; and the modular, or resources account (knowledge-in-pieces, p-prims, epistemological resources) is appropriate, when learners' cognitive formulations are context-dependent and labile (Hammer 2000; Taber 2008; Redish 2004; Scherr, 2007). Normally, experience in a discipline results in stable mental models; so, undergraduate students who take a non-majors physics course, will likely have context-dependent, or local coherence of concept and belief formations, with the possibility of cognitive crossover from their majors (Muis, et.al., 2006).

In the following sections I will first review the theoretical framework of epistemic framing as an emergent model of cognition, which accounts for both - the conceptual or knowledge-based, and epistemological dimensions of cognition, and next, I will review the empirical basis of the two interpretive techniques - epistemic framing analysis, and epistemic network analyses, which are based on the epistemic framing model of cognition, and are used to analyze learners' cognitive process, and therefore, sophistication of their thinking.

Epistemic Framing Theory

The epistemic framing theory (EFT) developed out of a perceived need for a unified framework which accounts for the conceptual as well as the epistemological components of cognition. While the 'coordination class and primitives' theory as conceived by diSessa and Sherin (1998), takes a structural approach, to model cognition, the epistemic framing theory takes a process-level binary approach. Per the epistemic framing theory, cognition functions at two levels – (i) knowledge structure, formed by the activation of patterns of associations (linkages or connections) between knowledge elements (reasoning p-prims, etc.) and resources (elements, schemas, mental models, etc.), and (ii) control structure, which conceptualizes the executive control of the activation process by expectations and epistemologies (Redish, 2004). Thus, EFT model is used to assess the cognitive process, by identifying learners'

epistemic frame – the set of knowledge and epistemological elements which are found to characterize learners’ thinking.

Knowledge Structure or Conceptual Resources

Tuminaro and Redish (2007) use the term knowledge element to describe ‘any knowledge (declarative or procedural) stored in long term memory that activates as a single unit or chunk. Such knowledge is said to be compiled, and a compiled knowledge element that appears irreducible to the user is called a conceptual resource (Redish, 2004). Knowledge is compiled by making new associations and suppressing old ones. Activation of one resource may result in the associated activation of other resources, and patterns of association may develop, linking different resources in different situations. The patterns of association individuals develop may help or hinder them in learning physics (Redish, 2004). ‘Difficulty’ arises when learners tend to activate inappropriate resources, and a ‘misconception’ results from a robust pattern of mis-association (Hammer, & Elby, 2003). Carefully designed instructional environments can lead to appropriate (expert like) patterns of association and provide substantial repetition and practice for knowledge compilation (Redish, 2004; Tuminaro, & Redish, 2007).

Control Structure or Epistemological Resources

Resources used by learners to decide ‘that they know something’ or ‘how, why, when they know something’ are called epistemic resources (Redish, 2004; Tuminaro, & Redish, 2007) or epistemological resources (Hammer, & Elby, 2003). “A student may “know” that a big car hitting a small car exerts a bigger force on the small car than the small car exerts on the big one because “the big one is stronger” (knowledge by p-prim)” (Redish, 2004, pp.30). Analogous to the associative structures of knowledge resources – schemas and models, associational patterns of epistemic resources are described as ‘epistemic games’ and ‘epistemic forms’ (Redish, 2004). Some of the common epistemic games played by students are ‘making lists’, ‘comparing entities’ ‘analyzing trends’ and conducting inquiries’ (Collins, & Ferguson 1993; Redish, 2004). A ‘list’ is the form completed by playing the list-making game, ‘an explanation’ or ‘a mathematical representation’ or ‘a schematic diagram’ may be epistemic forms completed or generated by playing an ‘inquiry game’ or a ‘meaning making’ game in a physics lab session (Tuminaro, & Redish, 2007).

Epistemic Frames and Framing

When presented with an environmental stimulus, a learner responds by parsing the input information (received from the physical world, culture, and social interactions) into pieces, attending to parts of it and ignoring the rest (Redish, 2004). Learner’s expectations and interpretations influence what parts of the sensory input get selected or attended to and perceived. The control process of choosing a set

of knowledge building tools stored in memory for use in the activity of knowledge construction, is called epistemological or epistemic framing, and the set of cognitive resources that a learner believes are appropriate to carry out a learning task are called epistemological frames or epistemic frames, or simply e-frames (Redish, 2004). The decision to choose a set of resources may be conscious or automatic, and is influenced by learner's contextual and pre-conceived beliefs or expectations about the learning task or activity (Tuminaro, & Redish, 2007).

The idea that a person's expectations – developed from prior life experiences, influence his/her interpretation of complex real-world scenarios which he/she experiences, and determine how he/she makes sense of the situation, has existed for a long time before it was utilized by researchers in PER. Bateson (1972), a psycho-analyst, first used it to analyze observations of monkeys' behavior of biting each other, calling the monkey's interpretation of the behavioral context, as monkey's frame, to explain how monkeys made sense of the context as calling for 'play' or 'fight'. Tannen (1993), a socio-linguist, used the concept to analyze a pediatrician's interactions with a child, the child's mother, and a tape recorder, which was recording the doctor's conversation with the mother about the child's illness in the presence of the child. She (Tannen, 1994) analyzed each of the three interactions of the doctor, as taking place in different frames, based on different vocabularies, language complexity, and voice tone.

Framing as a control process, and frame as a control/executive cognitive structure, were introduced in physics, to analyze the role of student's expectations in physics learning (Bing, 2008). If an undergraduate student's expectations (thinking) about what he/she is supposed to do (developed over many years of schooling), do not match the teacher's expectations of what the student is supposed to do, then even if students have the conceptual resources to answer a problem or complete a task, the resources may not be activated appropriately in that learning context, and may not get activated even after suggestions by the TA or the teacher (Redish, 2004, p.32; Scherr, & Hammer, 2009). Similarly, learners' epistemological beliefs about the nature of physics knowledge and learning, and what counts as valid sources of physics knowledge and physics learning, may also inhibit the activation of appropriate resources in a context (Scherr, and Hammer, 2009). 'Bridging Analogies' and 'Elby Pairs' are two techniques used by instructors, to address such problems (Redish, 2004).

Bridging analogies are lab activities, discussions, or other such means used to bridge the gap between target concept which learners commonly do not believe to be true, intuitively, and an anchoring example which they intuitively believe to be true (Redish, 2004). 'Elby Pairs' consist of pairs of questions which ask the same question in two different ways, such that students are likely to answer one question with a common misconception, while the second question is constructed to match student's intuition, and answered correctly (Redish, 2004). It is helpful in such instances of right or wrong intuitions about the

‘physics’ behind a concept or task, that instructors have knowledge of, and can identify p-prims and ‘resources’ which hinder or enable learning of the ‘physics’. The proposed ENA-based assessment mechanism will therefore be a significant step in this direction, since its primary goal is to identify and quantify learners’ conceptual and epistemological resources, in-context.

Resources. In examinations of learners’ epistemic frames in-context, the set of resources identified as constituting learners’ frames, and the cognitive structure of these resources have been ontologically classified in different ways by different researchers. While the basic elements (of knowledge, skills, beliefs, etc.) of these resources remain the same, the grain-size of the resources may vary between coarse and fine-grain sizes. The classification differentiates on the lines of (a) maturity level of the target audience - introductory versus upper level physics, majors versus non-majors, etc.) (Bing 2008, Tuminaro, 2004; Lising, and Elby, 2005); (b) the specific cognitive mechanism, and/process being examined (e.g. problem solving) (May, and Etkina, 2002; Scherr, and Hammer, 2009); and (c) the method of analysis chosen by the researcher – whether qualitative epistemic framing analysis or quantitative epistemic network analysis (Bodin, 2012). The following sections elaborate this ontological classification of resources through previous works in both categories. Epistemic framing analysis is a purely qualitative technique, and it is the first step in the process of epistemic network analysis; the next step being network analysis of the associations between the activated resources identified in the first step.

Epistemic Framing Analysis

Researchers have used a variety of observational methods to find evidence of framing. In each of the empirical examples discussed in this section, the method of interpretation has been qualitative, while observations have used techniques which cannot be used with large sample scenarios presented by a classroom or Lab session of an introductory university physics course. However, the qualitative coding process of learners’ tweets – method of observation proposed in this study, will be along the lines of epistemic framing analyses detailed below.

Lising and Elby (2005) chose to focus on only one facet of a single female student’s epistemology - how she selected and chose conceptual resources in learning introductory college physics, rather than a range of ideas about knowledge and learning. This enabled them to make in-depth arguments about causality, using one semester of videotaped data of two groups (with the female student being a member of one group), collected during three hours of interactive lectures - one hour of tutorial, and two hours of traditional lab; their written work; and six interviews with the female student, as she worked by herself on homework problems (Lising, & Elby, 2005).

Reasoning patterns in interviews were quantized using a coding scheme designed to pinpoint instances in which she used formal, classroom-taught reasoning versus ‘everyday’, and intuitive informal

reasoning; “when she was sense making versus just trying to remember or throwing out ideas with little thought; and when she attempted to reconcile different lines of reasoning” (Lising, & Elby, 2005, pp.378). The researchers then developed a hypothesis based on interview data, that the student’s inability to connect between formal (mathematical), conceptual, and informal reasoning was interfering with her learning. The ‘relationality’ or ‘connection’ barrier called epistemological barrier by Lising, & Elby (2005), evidenced in the student’s tendency to view formal and intuitive or everyday reasoning as separate, remained consistent across contexts.

Interview data was coded as ‘formal’ or ‘everyday (informal) reasoning; and 71 instances of informal reasoning, 13 instances of formal reasoning, and 13 of hybrid (integration of everyday and formal) reasoning were identified. Of the 36 instances of reconciliation (relational) opportunities, the student reconciled 13 out of 28 within-type reconciliations, and only one out of 8 between-type reconciliations, which was weak, because it was strongly prompted, suggesting that learners’ epistemologies can be explicitly attended to during instruction, and such ‘help’ enables the student to scale the epistemological barrier that prevents connections. The truth of the hypothesis was tested and verified on the student’s written work (Lising, & Elby, 2005).

May, and Etkina (2002) used weekly open-ended reflection instrument to assess learners’ epistemological preferences; and the questions asked were: “What did you learn in lab this week? How did you learn it? What did you learn in lecture and recitation this week? How did you learn it? What questions remained unclear? If you were the professor, what questions would you ask to determine if your students understood the material?” (pp.1250). The emergent coding scheme with respect to students’ responses on ‘what they learned’, had categories such as ‘formula’, ‘vocabulary’, ‘concept’, and ‘skill’; and for students’ responses to ‘how they learned’, some categories were, ‘learned by doing’, ‘predicted/tested/interpreted’, and ‘reasoned or derived in lecture’. Reliability of the coding scheme was ensured by checking that the codes applied by two independent coders, to the reports of four different students, matched perfectly for the type of code, and 90% of the time for number of indications of the code.

Scherr, and Hammer (2009) coded and used ‘behavioral clusters’ observed on videotaped data of students working collaboratively in inquiry settings as evidence, to identify the epistemic frames adopted by students as they reasoned with the inquiry tasks. Two researchers independently observed and time-stamped ‘shifts’ or ‘changes’ in behavior – “students’ vocal register, affect, grammar, gesture production, and body language”, to identify behavioral clusters – “hand motions including gestures, facial aspect, body position and/or movement, vocal register, gaze, and so on”, and coded them in real time without transcript (Scherr, & Hammer, 2009, pp.153). Inter-rater reliability was 95% on the identification of

cluster type, and 90% on the timing of the transitions, to 5 seconds' accuracy. Four behavior clusters identified were applied to three groups of four students, to correlate behavioral clusters with the substance of students' reasoning that accompanied the behavior, and infer the epistemological frames chosen by students.

Tuminaro (2004) used epistemic framing to analyze videotaped sessions of college students solving homework problems in groups, to model student thinking in the use of math in physics, and to analyze student errors in reasoning with calculators. 11 hours of data were selected from 60 hours of videotape, based on the criteria that 'they were rich in student thinking', and that there was 'explicit use of math', and transcribed. The video data was time-scaled and the scales roughly corresponded with epistemic frames (problem solving time-scale of about an hour), epistemic game (heuristic time scale of about 10 minutes for solving one problem), activation of knowledge elements (thought time scale of about one second for looking at an equation, or graph, and say something about it) (Tuminaro, 2004). Mathematical resources (Knowledge elements), epistemic games and epistemic frames were identified and coded using heuristic principles. Initial inter-rater reliability between two coders was 80%, which improved to 100% after discussions and refinement of codes (Tuminaro, 2004).

Bing (2008) conducted a similar study with upper level physics students, collecting 150 hours of videotaped data on students solving homework problems in groups, to model their thinking and reasoning with math in physics. Data was selected by searching the video database for 'arguments', 'debates', and 'misunderstandings', which were reasonably supposed to yield information on students' epistemic frames and shifts. Fifty snippets were identified based on the field notes taken by the researcher, alongside the videotaping, to conduct epistemic framing analysis, and identify clusters of framing, and to describe how students were interpreting the math used by them, or what type of warrants they used as arguments. The initial analyses were applied to other episodes, and refined for consistency, until all episodes (including the ones that were omitted initially) could be described using the model of student thinking developed (Bing, 2008).

The four clusters of epistemic frames identified were, "Calculation", "Physical Mapping", "Invoking Authority", and "Math Consistency" (Bing, 2008, pp. 43). In a 'Calculation frame', students typically relied on computational correctness; in a 'Physical mapping frame' they supported their arguments by pointing to the quality of fit between their mathematics and the physical situation at hand ; in a 'Invoking Authority' frame, students took the practical approach of taking some results for granted without going to the level of first principles; and in a 'Math consistency' frame, students cited consistency of given problem with a more familiar use of the mathematical idea in play (Bing, 2008).

This section discussed examples from PER which targeted the same population being targeted through this study – students of introductory university physics courses. Its purpose was two-fold – (i) to elaborate the technique of framing – data collection and analyses, which is also the first step in epistemic network analysis (ii) to compensate for the void in ENA literature, of example studies investigating similar contexts and similar audience levels in physics, targeted by this study. It should be noted, that the data collection techniques used in these studies, were videotape (Tuminaro, 2004; Bing, 2008), interviews (Lising, and Elby, 2005), observations (Scherr, and Hammer, 2009), and/reflection instruments (May, and Etkina, 2002), which generated non-digital data, requiring cumbersome pre-processing prior to analyses. To obtain data on learners’ thinking process, rich with learners’ own articulation, and usage of canonical tools of the physics community, the ‘reflection instrument’, which collected self-reported data on learners’ thinking (May, and Etkina, 2002), presents a good fit. While technology can be integrated to adapt any of the above data collection techniques, to yield digitized data, the data obtained from other three tools, is devoid of learners’ articulation of their ‘inner speech’ in the words of Vygotsky. Sophistication of thinking, in this study, was determined from learners’ own articulations by way of descriptions, justifications and/ explanations of concepts and/findings from experiments. The following section will elaborate on the specific observational and interpretive techniques used in epistemic network analyses.

Epistemic Network Analysis

The ontological classification of resources with this method of interpretation is based on the elements of knowledge, skills, beliefs, strategies, ways of making decisions (epistemological rules), values, and/identities which typically characterize complex thinking in a domain, and in a specific context (Shaffer, et. Al., 2009). Expertise in these contexts can be modeled as a network of connections between specific understandings articulated through discourse. Epistemic network analysis (ENA), allows a way to quantify the connections between epistemic frame elements, and thus assess their development (Shaffer et al., 2009; Shaffer & Graesser, 2010). It focuses on patterns of relations or connections rather than individual elements which frame the discourse observed, and treats epistemic frames as cognitive networks, in which the different frame elements become nodes, while the patterns of connections constitute the links between these nodes (Svarovsky, 2011; Arastoopour, et. Al., 2016).

Bodin (2012) investigated university physics students’ cognitive representations, in terms of the epistemic frames associated with solving a computational physics problem. Construct of epistemic framing was chosen to describe the organization of knowledge, skills, and beliefs, which together were classified as the basic resources – epistemic elements. Construct of Epistemic network was used to refer to the network of cognitive organization corresponding to an epistemic frame, consisting of nodes,

representing epistemic elements identified from the interview data of six students, on two occasions. The six students were chosen by purposeful sampling of 36 participants enrolled in the course, by administering a standardized epistemological beliefs survey, the Colorado Learning Attitudes about Science Survey (CLASS), to generate a sample consisting of two students each from the categories of novice, intermediate, and expert (Bodin, 2012). Interviews of 25-15 minutes were conducted before and after students completed the assignment, and questions were designed to get information about what knowledge, beliefs and skills were used and how they were used in the problem-solving process (Bodin, 2012). Students were asked to conceptually describe the task, the strategies they planned to use, the difficulties they anticipated and how they would resolve them, and their expectations about gaining new knowledge and skills. Post interview followed up on the same questions to determine discrepancies between expectations and actual experience; emphasis was on how students troubleshooted their solutions and overcame difficulties (Bodin, 2012).

Transcripts were thematically coded by breaking them down into statements, and analyzed as knowledge or beliefs (Bodin, 2012). The epistemic elements coded as knowledge were further categorized as – physics, math, programming, modeling; and the beliefs as – expectancy or efficacy (self-perceptions of competence), value (reason for engaging in a task), and attributional (beliefs about how to attribute their achievement outcomes, e.g., effort, luck, ability, difficulty). An exploratory, iterative coding process resulted in the identification of a third category – resources (used to perform the assignment, e.g., literature, teacher, peers, feedback from visualization), and a total of 77 unique epistemic elements (Bodin, 2012). Coding scheme was unique to the context, and was expected to reflect the expert's knowledge base. From the coded transcripts, epistemic networks were modeled by treating semantic adjacency between elements as a link or connection of value 1, the weight of the link being determined by the number of times it was found in a transcript. All links were considered bidirectional, and nodes with no links were removed, resulting in symmetric matrices to represent the relations. Each student generated two networks, and networks of all students were added, to compare common epistemic framings before and after the assignment (Bodin, 2012).

The networks were visualized using the map equation algorithm (which optimizes the path of a random walker in a network and describe how information flows in the network), and corresponding visualization tools – map generator used to show structure of networks and alluvial generator to show changes between networks (Bodin, 2012). The optimization process clusters groups of nodes into modules where information is concentrated, and nodes are linked to a higher degree. The stability of modular structure was examined by performing significance analysis, by resampling bootstrap networks. PageRank measure was computed to measure the relative importance of the module – the amount of focus

students put on epistemic elements in that module. It was found that the patterns of relations between elements increased in complexity, suggesting a more coherent view of the task change after the task; students' use of physics concepts after the task increased by 100%; their PageRank measure increased from 17% to 24%; and use of Math in students' framing decreased after task. Belief elements were present on both occasions to the same amount, however their interaction with other epistemic elements were found to change after the task (Bodin, 2012).

Arastoopour, Shaffer, Swiecki, Ruis, & Chesler (2016) examined the final epistemic frames of 45 (13 female, 32 male) undergraduate engineering students enrolled in an introductory engineering course at a large public university, after they participated in a virtual internship, Nephrotex, as interns at a fictional biomedical engineering design company. Purpose was to examine the effects of team work and engineering design on students' motivation to continue in engineering. Students worked collaboratively (in groups of 3) on designing dialyzers for hemodialysis machines, using a customized simulation of dialysis membrane design, to develop and test their devices, while addressing the concerns of company's internal consultants. Research and design activities and group interactions took place through a web platform and consisted of – examining research reports on experiments with different materials and processes, developing hypothesis (individually and then in groups), testing hypotheses using the simulated engineering drawing tool and analyzing results. At the end, students presented their prototype-design to their peers and instructors, explaining and defending their design. Data consisted of pre-post survey on perceptions about and motivation to pursue engineering, and chat discourse (Arastoopour, et.al., 2016).

Principal component analysis was used to analyze students' scores on a four-point Likert-scale rating of a survey instrument comprising 20 questions on perceptions and commitment to the field of engineering (Arastoopour, et.al., 2016). ENA was used to measure the development of connections between the epistemic elements classified as knowledge, skills, identity, epistemology, value, and identity. Coding was done by segmenting all chat discourse into utterances, which are instances of student sending a single chat message. Twenty-one codes were developed from ABET criteria for undergraduate engineering program outcomes and using epistemic framing theory as a guide for professional practices. Links or connections between frame elements were defined as co-occurrences of the codes within each utterance, and each utterance was coded for connections by treating a single occurrence as 1, and no occurrence as 0 (Arastoopour, et.al., 2016). Data was segmented into stanzas defined by class sessions. Each coded utterance was modeled by the ENA Web tool as a binary vector, each coded utterance vector was converted into an adjacency matrix, represented by an adjacency vector, summed into a cumulative adjacency vector for each participant in each segment, for each stanza, normalized, and singular value

decomposition performed to interpret the connection pattern (structure) and contribution of each frame element within the components examined. Results showed that women demonstrated significant increase in positive views on engineering; and that participants who made more connections between engineering design and other epistemic engineering elements showed more positive changes towards career in engineering, than those who made more connections with collaboration (Arastoopour, et.al., 2016).

Svarovsky (2011) explored the development of complex thinking in ten middle school girls in the context of participating in Digital Zoo, a four-week summer program which provides an authentic learning environment to engage in authentic engineering activities designed to enable connections between engineering knowledge and skills, with engineering ways of thinking. Participants role played as bio-mechanical engineers designing solutions to a client-based project which required them to design character prototypes for an animated film, using SodaConstructor, an online spring-mass modeling system, to engage in iterative design-build-test cycles to create their designs. Players maintained a design notebook, working under the guidance of an undergraduate design advisor to develop designs, and provided updates and received feedback from clients (engineering graduate students) on a regular basis. Data consisted of pre-post and follow-up clinical (3 months after the end of game) interviews, and in-situ data – copies of participant-produced work, recordings and field notes of design meetings and conversations, and occasional videos and photos (Svarovsky, 2011).

Interview data was analyzed for determining learning outcomes, while ENA was used to analyze the learning process through analysis of in-situ data. In-situ data for the first two weeks (total of 30 hours) was converted into design histories of individual players (Svarovsky, 2011). Digital histories and interview transcripts were coded for five elements of epistemic framing – knowledge (use of professional terms of art and science vocabulary), skills (references to engineering competencies e.g. brainstorming, comparing alternatives, etc.), identity (references to role played by players or engineers), values (adhering to client need, creating an optimized design, etc.), and epistemology (ruling out a design because it was too costly, evaluating tradeoffs when making a decision, etc.) (Svarovsky, 2011). The interview code frequencies were tallied using paired-sample t-tests for pre-post means of frame elements to determine learning gains. ENA was conducted on the design histories, by segmenting them, first by week; further segmenting by focus of activity (client-focused, or non-client focused), and then by type of reflection (design meeting or design notebook), to get a total of eight segments per player (Svarovsky, 2011). Patterns of increasing or decreasing relative centrality of frame elements across different activity structures were explored and identified. Relative centralities of knowledge and skills followed similar trajectories, and remained high throughout the game; but reached a low point during the segment (Week1, Client project, design meeting); values and epistemology seemed to become central during client-focused

activity and notebook based reflection during Week1, identity during non-client work and design meeting of week1; and all three decreased in centrality in week 2 (Svarovsky, 2011).

Conclusions

In the empirical works discussed above, excluding the study in computational physics (Bodin, 2012), digital artefacts generated during task performance, were used as data collection tools, and ENA Web Tool, was used for network analyses of the associations between code categories discovered in the qualitative data collected, using co-occurrence of codes to define an association. Although Bodin (2012) used contextual interviews for data collection, network analysis was similarly conducted, of the associations between code categories discovered in the data, using semantic adjacency to define an association, and the map equation and alluvial generator software developed by researchers at Department of Physics, Umea University, Sweden, for network analyses. The ontological classifications of cognitive elements in these studies, have been similar, but distinct from the epistemic framing analyses, in that the focus is to map cognitive networks of learners in-context, rather than to identify the epistemological resources – beliefs, expectations, values, identities. Thus, although ENA is framed in the theoretical lens of epistemic framing theory, the focus is on network analyses of associations between cognitive elements, and the process of developing of the coding scheme is driven by researchers' questions, purpose, and type of data (Shaffer, 2017). Therefore, the code categories may, or may not, explicitly include a variety of epistemological resources.

Students' difficulties in physics have often been attributed to their difficulty in bringing to bear 'the math knowledge' they are expected to possess, since they are required to have completed the requisite number of math courses – algebra, precalculus, and/geometry. These issues have been analyzed extensively in PER and attributed to learners' inability to relate a given learning context in physics, to their 'repertoire of math knowledge' (Bing, 2008; Tuminaro, 2004). In other words, a lack of connectedness, or connectivity, or connections between relevant mathematical equations, graphical relationships, and physical quantities under study, is the root cause of the issue, which is often termed in ENA vocabulary, as a 'framing' issue. Students are unable to correctly choose or determine which 'mathematical resources' are relevant or count as appropriate in a problem-solving situation in physics. In Epistemic framing analyses, learners' warrants or justifications or reasons for choosing a resource, are identified, and classified, or coded (Redish, 2004), while ENA does not require explication of warrants – it can be conducted by simply classifying or coding learners' responses, in an emergent, but theory-driven way, which seeks to differentiate the levels of sophistication of connections between the cognitive elements identified in learners' thinking (Shaffer, 2017).

In this study, the two approaches were used in a syncretic manner, meaning that warrants were examined, and coded, in responses to questions, which required learners to justify, and in other instances, the epistemological stance was coded, by interpreting the source of knowledge element, since learners seemed to have counted the source as valid to the context, consciously, or unconsciously. So, Vygotsky's classification (Vygotsky, 1962; Otero, 2004) was extended, to further differentiate the types of sources, learners seemed to have drawn upon, to construct their response. When canonical knowledge and tools were used, the corresponding knowledge element was classified as canonical (analog of scientific concept), and the spontaneous, or intuitive category was elaborated as concrete-intuitive and abstract-intuitive knowledge elements, to differentiate between knowledge elements which appeared to be (a) vague relationships between physical quantities related to the concept, constructed by the learner, that were canonically or scientifically inappropriate, and (b) relationships constructed in terms of concrete entities which embodied the physical quantities, or relationships, which were partially correct, but seemed to have been constructed from learners' interpretations of course readings. Category (a) was termed as abstract-intuitive, and category (b) was termed as concrete-intuitive.

The other adaptation, (or extension) of ENA implemented in this study, was the addition of an exclusively quantitative analytical step, in the assessment process, to quantize and operationalize the construct of sophistication of thinking. While ENA analyses discussed above, have utilized either the limited quantitative aggregation and statistical description afforded by the ENA Web Tool, or the Map Equation Generator, I have chosen to integrate another software package (RStudio), to augment the network analyses, with quantitative analyses, to enhance the discriminatory power of the assessment, and to enable its usage as a classroom formative assessment in inquiry settings in university physics, somewhat like Clickers. Further details are provided in the following chapter, which reports the methodology adopted to implement this study, along with its rationale, design, data collection, analysis, and validation methods chosen.

CHAPTER 3 METHODOLOGY

Inquiry learning environments which integrate technologies such as Simulations, Spreadsheet, and various computer-aided measuring instruments with graphical output displays, are commonly employed in the design of introductory university physics courses (Meltzer, and Thornton, 2012). Activities are structured to have learners explain-predict-describe-analyze-justify the findings of their inquiry, problems and/experiments, with the goal of promoting the development of appropriate connections between concepts, their mathematical, graphical, and schematic representations, and reality (Laws, Sokoloff, and Thornton, 1999). However, existing assessment instruments typically used in these environments, e.g. in-class and home-work Quizzes, and Lab Reports/Logs, fail to capture the ‘various connections’ in learners’ thinking, either because the data collected is quantitative, or the rich qualitative data collected is not analyzed to uncover the connections in learners’ thinking, as in the case of Lab Reports (NRC, 2001; Larkin, 2011). While Lab reports can provide valuable information about the thinking process, I contend that the raw qualitative data contained in them remains mostly untapped, since rubrics used to assess them, merely look at the ‘accuracy of measurements’, and ‘completeness of tasks assigned’, rather than analyzing student reasoning (McDermott, 2001).

Epistemic Network Analysis has emerged as a reliable analytical tool to model connections between various elements of knowledge, skills, and epistemologies which are found to co-occur in learners’ thinking in-context and has been used effectively by researchers in the context of virtual internships and games in the domain of Engineering, using conversational data such as chat board messages (Arastoopour, et.al., 2012). However, it has not been tested in the context of inquiry activities in physics, which are conducted in face-to-face Lab settings, although its usage could provide instructionally valuable information about learners’ thinking, as discussed in the previous two chapters. While inquiry lab activities generally require peer-to-peer interactions (Furtak, et.al., 2012), which can be potentially good sources of conversational data, the processing of such oral communication data can be cumbersome. Social media communication platforms can offer good alternatives, especially since the target population of learners prefer these modes of expression and communication (Chen, and Chen, 2012). This study explored the potential of one such preferred medium – Twitter.

As discussed in Chapter 1, the purpose of this study was to leverage four technologies – Twitter, Spreadsheet, ENA Web Tool, and RStudio, to develop and test the validity and reliability of an assessment framework, which utilized the lens of epistemic network analysis (ENA), to assess the sophistication of learners’ thinking, during collaborative inquiry activities conducted in studio lab setting of a non-Majors university physics course. The overarching question which framed this study was:

“What is the validity and reliability of using an ENA-based assessment implemented by interfacing the technologies - Twitter, Spreadsheet, ENA Web Tool 4.0, and RStudio, to assess the sophistication of learners’ thinking in the context of four inquiry experiments conducted during a studio lab session in an undergraduate physics course?”

In this chapter, I will describe the methodology that was used to achieve the purpose of the study, as stated above. Included in this chapter is a visual representation of the research design, in addition to a rationale for choosing the design. Explanations are provided for how quantitative and qualitative methods worked together to achieve the purpose of the study, in addition to the role of the researcher, sample selection, data collection, data analyses, and instrumentation. Trustworthiness of the findings will be discussed, and the chapter will conclude with a summary of the methodology used for this study.

Participants and Background

The participants for this study were drawn from a population of undergraduate students enrolled full-time at The University of Tennessee, Knoxville, taking the non-majors undergraduate physics course ‘Elements of physics’ (PHY222) during the Fall semester of 2017. The average undergraduate enrollment at the university is 20,000 students, with an equal number of male and female students, but the student body is predominantly White by demography. The course satisfies four credits of the general education requirement for graduation, under the natural sciences category. Students who take this course are typically from pre-medical, pre-dental, pre-pharmacy and pre-veterinary programs, in their second year of college, and take Elements of Physics ‘PHYS 221’ (a pre-requisite for 222), an introductory course in ‘Mechanics, Heat and Sound’ in the previous semester. Typically, students have some background in Math, not necessarily in Calculus, since PHYS 221 requires them to have this knowledge. The study was administered to two sections of PHY222 – Section 001, and Section 003, which met on the same day, but at different times of the day – morning and afternoon respectively, for the studio lab sessions. The total number of participants was 38.

Setting of Study

The course covers ‘basic physical principles and applications in Electricity, Magnetism, Optics and Modern physics, required in pre-medical, pre-dental, pre-pharmacy and pre-veterinary programs’. The course is typically taught in a lecture-lab format, which consists of a large lecture (5 sections together in Fall and Spring) session of one hour fifteen minutes’ duration, and a studio laboratory session (45 students per session in Fall and Spring) of two hours and forty-five minutes’ duration, per week. The large lecture and studio sessions are inquiry-based, and make use of PowerPoint, clicker questions, simulations, spreadsheets, applets, and hands-on inquiry activities designed to provide learners opportunities to use, interpret and translate between multiple representations of the physical concept

under inquiry, and to predict-observe-explain. They are thus actively engaged in reasoning and connecting between concepts, their graphical, mathematical, and/ schematic representations, and the physical reality. Students are expected to read instructional materials, and complete homework quizzes prior to the lecture as well as the studio sessions. All curricular material is made available online through Canvas – the university’s course management system

The studio sessions are directed by two teaching assistants, one of who starts with a PowerPoint presentation reviewing the main ideas related to the topic. A home work Quiz administered online, prior to the lab session, and a Lab Quiz, administered during the studio session, but before the inquiry activities, enable the TAs to get an idea of the learner’s level of understanding of the topics they are expected to enquire. A brief overview of the actual activities structured into Studio 3 is given below. The activities are timed, and the students work in groups of 5 or six students at a round-table, having a computer for the students to access learning materials. The computer is connected wirelessly to a printer in the lab, for students to print out their worksheets, before or after completing the inquiry. Students are also expected to work on their personal laptop computers. The TAs move around, listening in to the group discussions and providing ‘help’ when asked for.

Contexts of inquiry. The studio session targeted for this study was ‘Studio 4’, which covered the topic of ‘electric current and electric circuits’. The session included PowerPoint presentation and a lab Quiz, followed by four hands on experiments: (i) Investigating resistance of human body (Experiment 1) (ii) Measuring Capacitance (Experiment 2) (iii) Investigating Ohm’s law (Experiment 3) (iv) Investigating RC circuits (Experiment 4). Students completed inquiry tasks in groups of 3, and typed in their answers to specific ‘predict-observe-explain questions, on a Microsoft document, which was printed out, along with graphical outputs generated during the experiments, as a report or log of their activities (Labs), and submitted to the TAs. The reports were graded by the TAs for 100 points. The specific inquiry tasks, and questions, for each experiment, are appended in Appendices 1a, 1b, 1c, and 1d.

Research Questions

The specific research questions (RQs) which informed the design of this study were:

RQ1). What themes of ‘epistemic framing’ can be discerned from qualitative analyses of learners’ Twitter responses, to question prompts sent to them before and after each activity, about the physical concepts being investigated through each of the four inquiry experiments? How do the pre-and post-activity epistemic frames compare qualitatively? (QUAL)

RQ2). (a) What mean patterns of connections do the Network Graphs and Matrix Data reveal in learners’ thinking before and after each of the four inquiry experiments (activities)? (b) Do the pre-and post-activity patterns differ significantly by (i) Section of course and (ii) Levels of (a) Epistemic beliefs

Score (b) Lab Quiz Score (c) Lab Score, (d) Homework Quiz Score (Attempt1), and (e)Test1 Scores? (QUAL+QUAN)

RQ3). What levels of sophistication were identified in learners' thinking, through exploratory statistical analyses, of scores of node-centrality and weight density of connections? How do the measures of centrality and connection density of learners' Epistemic Networks relate to (i) Lab Score (ii) Quiz scores, (iii) Test1score, (iv) A3 HW Quiz score, and (iv) EBAPS score? (QUAN)

RQ4). (a) How do the inquiry activities compare with respect to their effect on the sophistication of learners' thinking? (b) Does the proposed ENA-based assessment framework provide a valid and reliable tool to assess the sophistication of students' thinking in inquiry contexts in undergraduate physics courses? (MM)

Research Design

As discussed in the previous chapters, and the paragraphs above, the purpose of this study was to (a) develop a technology-enabled, ENA-based assessment, suitable for large-sample assessment of gains, in sophistication of connections in learners' thinking, during inquiry experiments in physics, and (b) test the validity and reliability of this assessment. Thus, there were two design goals – developing the assessment, and testing its validity and reliability. Each of these two goals contained multiple objectives, and are therefore discussed separately, to present the rationale for adopting a single-study, multiple phase approach to mixed methods research design.

Design Goal 1: Development of Assessment

To assess development of connections in learners' thinking, during inquiry experiments, using ENA as described in chapter 2, the following steps would be involved –

- (i) gathering qualitative observations of learners' thinking, or conversations,
- (ii) qualitative coding of observations, to identify categories of sophistication of knowledge,
- (iii) converting the co-occurrences of codes into adjacency matrices, and networks of nodes and edges,
- (iv) analyzing the networks.

To compute gains in sophistication, the network and matrix data generated by ENA, would have to be further analyzed quantitatively. Thus, to implement the ENA-based assessment and compute sophistication gains, the following processes would have to be executed sequentially - qualitative data collection, qualitative analyses, network analyses, and quantitative analyses; each process essentially

building on the previous process, to generate enhanced understanding of ‘sophistication of learners’ thinking’ (Figure 3.1).

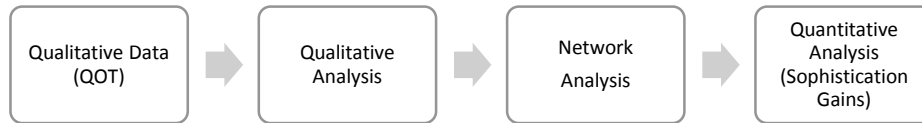


Figure 3.1. ENA-based Assessment Framework.

To enable large -sample implementation of the assessment as required by inquiry contexts in physics, technologies compatible with each other, would have to be seamlessly interfaced. While ENA Web Tool 4.0, a software package developed by developers of the ENA technique, was available to conduct network analyses, there was no precedence found in literature, of technology usage to (a) collect large-sample qualitative observations of thinking (QOT) in inquiry contexts in physics, and (b) conduct quantitative analyses of network and matrix data generated by ENA Web Tool. This study explored the use of Twitter for collecting large sample QOT, and RStudio for computing sophistication gains quantitatively. Spreadsheet was used for qualitative analyses, because of its compatibility with Twitter, and with ENA Web Tool. The rationale for technology choice is discussed below in data collection and analysis sections respectively. The technology-enabled assessment is illustrated by figure 3.2.

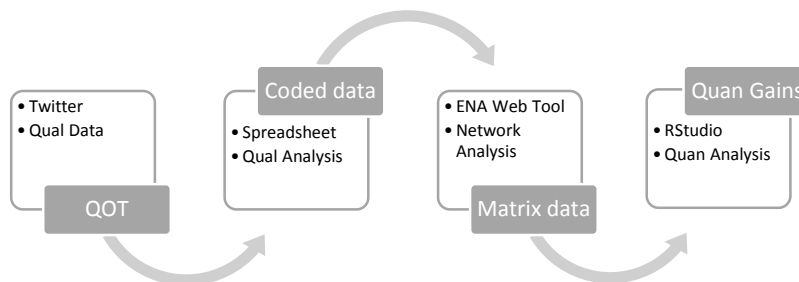


Figure 3.2. Technology-enabled Assessment.

Exploratory design. The development of assessment was essentially exploratory because there was no precedence found in literature of the following – use of the framework proposed in Figure 3.1, to assess sophistication gains in physics; use of ENA in inquiry contexts in physics; use of Twitter to collect large sample qualitative observations of learners’ thinking in these contexts; and use of RStudio for computing sophistication gains by conducting quantitative analyses of network and matrix data.

Design Goal 2: Testing Validity and Reliability of Proposed Assessment

Validity refers to how accurately the assessment instrument measures the item of interest – ‘sophistication of thinking’ in the case of proposed assessment, and reliability refers to whether the instrument gives the same results each time it is used in the same setting with the same subjects (Sullivan, 2011). Establishing validity requires the construction of an evidence-based argument regarding how well the tool measured what it was supposed to measure, and there are different ways by which the evidence can be generated (Sullivan, 2011; Adams, 2010). Similarly, reliability can also be estimated in different ways (Sullivan, 2011). The choice of methods depends on the assessment and the context (Trochim, 2006). Typically, when a construct’s validity is being examined, as in the case of ‘sophistication of thinking’, construct validity is established by testing for one or more of the following (Trochim, 2006):

(i) Predictive validity, or the ability of the proposed assessment measures to predict performance it should theoretically be able to predict. It can be confirmed by a high correlation between the two constructs.

(ii) Concurrent validity, or the ability of proposed assessment measures, to distinguish between groups it should theoretically be able to distinguish. It can be confirmed by administering the proposed assessment to different groups which can theorized to show differences in sophistication.

(iii) Convergent validity, or the degree to which the results of proposed assessment converge with results of similar assessments, they theoretically should. It can be confirmed by high correlation between the results of both assessments.

(iv) Discriminant validity, or the degree to which the operationalization of proposed assessment, diverges from other assessments it should theoretically be dissimilar to. It can be confirmed by low correlation between the results of both assessments.

To examine validity, it was therefore necessary to identify assessments which could be theorized to be similar and/dissimilar from the proposed assessment measures. Since the premise of the study, as described in previous two chapters, was that existing course assessments fail to capture connections in learners’ thinking, low degree of correlation could be expected, theoretically, between the proposed measures and scores on existing assessments. Thus, theoretically, discriminant validity could be expected to be established, and convergent validity could be expected to be disproved, through correlational analyses between proposed assessment measures and scores on existing assessments. Concurrent validity could be examined by grouping participants on scores on existing assessments, and then examining whether there were statistically significant differences in network connections of these groups. Predictive validity could be examined through correlational analyses between proposed assessment and existing

assessments which could be theorized to measure constructs, like the construct of ‘sophistication of thinking’.

Sophistication of thinking, as discussed in chapter 2, has been defined and assessed in physics, in terms of performance on problem solving tasks, inquiry tasks, conceptual questions, and/ epistemological survey instruments. The pre-lab home-work quizzes, in-lab quizzes, and mid-term test, of PHYS 222, which can be said to take place in close proximity of time, with reference to administration of the proposed assessment, consist of conceptual and problem-solving tasks. The Labs themselves consist of descriptive and reasoning questions related to the experiments conducted. Scores on these assessments are thus, considered to be indicators of level of sophistication of learners’ thinking, and can be used to examine validity of proposed assessment. However, none of these assessments measure learner’s epistemic beliefs, which, per the theoretical model of cognition, used to develop the proposed assessment, influence learners’ responses on all assessments. Therefore, a separate survey could be administered to include assessment of sophistication of beliefs, for validation purposes. Details of the assessments and procedures are explained in a later section.

Reliability. Four reliability measures are typically used to examine reliability or consistency, or repeatability of an assessment: (i) Inter-rater reliability (ii) Parallel forms reliability (iii) Test Retest reliability, and (iv) Internal consistency reliability (Trochim, 2006). Parallel-forms and test-retest are administered when the construct examined can be theoretically assumed to remain unchanged. They were unsuitable in the context of this study, since connections in thinking are theoretically assumed to be context-sensitive. Thus, inter-rater reliability and Cronbach’s alpha (measure of internal consistency reliability) were chosen for designing the tests of reliability. Details will be discussed in a later section, following the data analyses.

Exploratory design. The validity testing of proposed assessment was also exploratory in nature, because the ENA-based assessment was designed to measure sophistication of thinking by explicitly measuring the strength and type of connections in learners’ thinking, while existing Course assessments, and the beliefs survey, measure sophistication in terms of demonstrated occurrence of specific elements of knowledge, skills, and/ beliefs. Thus, validation can, at best be explored, through correlational analyses with existing assessments, each of which only measures occurrences of specific elements, unlike the ENA-based assessment, which targets co-occurrences. Lack of assessments which measure the same construct – cognitive connections, makes it necessary to examine validity by including a range of assessments, which are broadly treated as indicators of sophistication of thinking.

Rationale for Using Mixed Methods

The following rationales have been identified in mixed methods literature, as prime reasons for adopting a mixed methods design: (i) Development or building of one method (either qualitative or quantitative) from the results of the other, (ii) Complementarity, or mixing methods to complement, enhance, illustrate, or clarify results (Caracelli, & Greene, 1993), (iii) Instrument fidelity, or mixing methods to assess the appropriateness and/or utility of existing instruments, by creating new instruments, and/or monitoring performance of human instruments; (iv) Significance enhancement, or augmenting the interpretations of one method by using another method (Collins, Onwuegbuzie, & Sutton, 2006), (v) Triangulation, or seeking corroboration, convergence, and correspondence by mixing methods; (vi) Initiation, or recasting results from one method type, with results from the other method; (vii) Expansion, or extending the breadth and range of inquiry by using different method types for different inquiry components (Caracelli, & Greene, 1993).

The research design goals of this study, as described in the previous subsections, can be re-stated as: (a) building on the results of qualitative analyses, by conducting network analyses, and building on the results of network analyses by conducting quantitative analyses, to generate enhanced understanding of ‘sophistication of learners’ thinking’ (b) assessing the validity and reliability of ENA-based assessment, by comparing its performance with existing instruments. These goals align with the rationales of ‘development’ (Caracelli, & Greene, 1993), instrument fidelity, and significance enhancement (Collins, Onwuegbuzie, & Sutton, 2006), described in the above paragraph; thereby justifying the use of a mixed methods design for this study.

Typology of mixed methods design approaches. The design was based on the fundamental definition of mixed methods research as ‘the collection, analysis, and integration of quantitative and qualitative data’, to understand a research problem (Creswell & Plano Clark, 2007). Creswell and Plano-Clark (2011) have identified the following broad categories of mixed methods design: Convergent Parallel, Explanatory Sequential, Exploratory Sequential, Embedded design, Transformative design, and Multiphase Design, based on four factors. (i) The ‘level of interaction of two strands – qualitative and quantitative, or the extent to which both are kept distinct or interacted. (ii) The ‘priority of methods’ for answering the research questions - whether both have equal importance (QUAL-QUAN), or qualitative methods have priority (QUAL-quan) or quantitative has priority (QUAN-qual). (iii) The timing or sequence of data collection and interpretation – whether they are concurrent, sequential or multi-phase. (iv) The mixing or integration strategies adopted - whether at the interpretation stage, the data analysis stage, data collection stage (embedded), or the design stage (theoretical framework) (Terrel, 2012; Creswell, 2015).

As discussed in the design needs of this study above, quantitative and qualitative methods were needed to be mixed at various stages, and multiple points of interface - at the data interpretation stage, at the data analysis stage, and at the research design stage. Quantitative data in the form of scores on various assessments was required to be collected at various points in time – for instance, the beliefs survey had to be administered first, to inform the qualitative coding process, as explained in the ‘procedures’ section. Network analyses by design, include qualitative and quantitative interpretations. Thus, the design was necessarily an ‘interacted design’. Its implementation called for a combination of qualitative and quantitative methods, which do not neatly fall under the common typologies of concurrent, sequential and designs. While the quantitative methods seemed to have priority, the data collection is predominantly qualitative, and the research purpose is grounded in the belief that the construct under investigation – learners’ thinking cannot be assessed or understood completely by either method standing alone. Thus, philosophically, the research gives equal priority to both methods, incorporating a combination of quantitative and qualitative methods at various stages (collection, analysis, interpretation) of the study as depicted in Figure 3.3. The basic design approach, was thus, ‘Interacted QUAL-QUAN’ wherein the qualitative and quantitative strands interacted throughout the study, as opposed to being purely concurrent or sequential, and both strands had equal priority.

Single study multi-phase, Interacted, QUAL-QUAN mixed methods design. The basic rationale for a multi-phase design is that the researcher examines a problem, through a connected, sequential implementation of qualitative and quantitative strands of research, such that the preceding phase informs the implementation of the next phase (Creswell, & Clark, 2007). The implementation of an ENA-based assessment, and its validation would require such implementation as explained in the ‘design goals’ sections, and as was illustrated in Figure 3.1 above. Thus, this study adopted the single-study multi-phase variant of multi-phase design, which combines both sequential, parallel, and/nested designs at various stages of a single research study (Creswell, 2011). The other two variants are – multi-phase program development/evaluation projects which require exploration, e.g. feasibility studies; and multi-level studies in which different methods and phases are used to examine different levels within a system, e.g. School, district, teachers, classrooms, students, etc. (Creswell, & Clark, 2007) As previously explained, the study is also exploratory in nature.

Design of the study (Figure 3.3). In this Dissertation study, each phase of the study refers to a distinct point in time, of data collection, analysis and/integration. While detailed procedures are outlined in a later section, a brief overview is presented here. In phase 1, quantitative data was collected in week 2, using EBAPS (Epistemological Beliefs about Physical Sciences) survey, and analyzed. In Phase 2, qualitative Twitter data was collected in week 4, and analyzed qualitatively, using the results of EBAPS

for random selection of Twitter responses of students from low, medium and high scorers, to achieve inter-rater agreement. In phase 3, quantitative data – scores of four existing Course assessments was collected after the mid-terms were administered, analyzed to group students by their scores on each assessment, and the results informed the quantitative component of network analyses – conducting t-tests. In phase 4, quantitative analyses on network data; correlational analyses between the computed measures of sophistication, and scores on existing instruments for establishing validity and reliability; and integration of results of qualitative, network and quantitative analyses from phase 2, 3, 4, were conducted.

Phase 2 and 3 are conducted sequentially, such that network analyses of Phase 3 are conducted by (a) transforming the results of qualitative analyses of phase 2, and (b) using the results of quantitative analyses embedded in phase 3. Such an approach can be viewed as a combination of two basic approaches – exploratory sequential design, and embedded design. Part (a) is like ‘instrument development’ variant of exploratory sequential design in which qualitative findings inform the development of quantitative measures, to answer different research questions (Creswell 2006). In this study, the qualitative findings were expected to answer RQ1 and were used to derive the co-occurrences of codes for conducting network analyses to answer RQ2a. Part (b) is like a variant of embedded design in which a supplemental dataset is embedded within a larger design to answer different research questions (Creswell 2006). In this study, quiz score, labs score, test score, and EBAPS score data were embedded in Phase 3 of study, to answer RQ 2b. It is important to note that the standard exploratory sequential design typically has two separate sets of data, whereas this study utilized the same sample to obtain the quantitative metrics (co-occurrences) from the qualitative data.

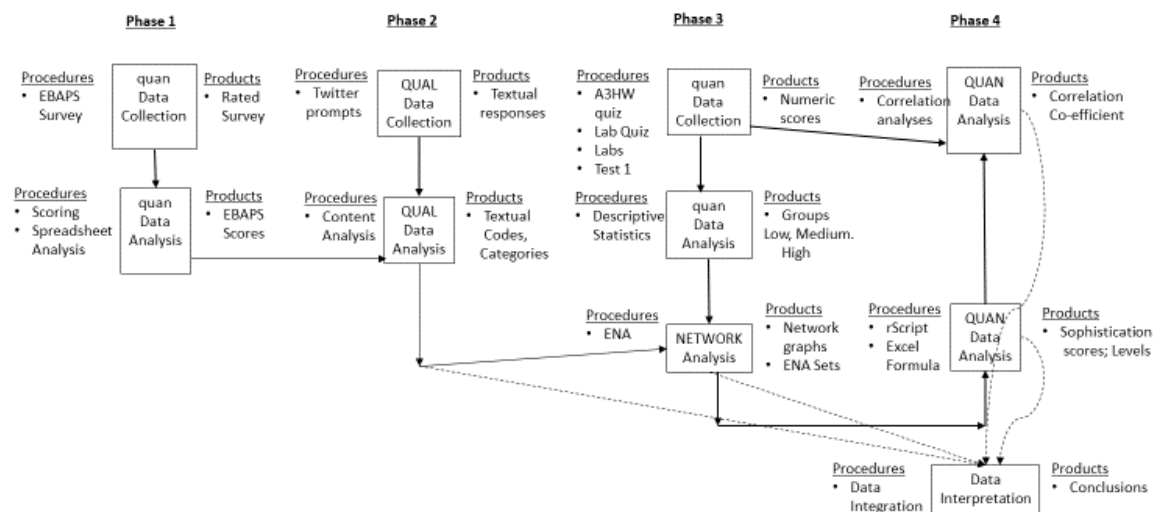


Figure 3.3 . Diagram of Research Design: Single-study, multi-phase, interacted, QUAL-QUAN mixed methods design.

Procedures

The proposed study complied with all ethical guidelines and completed the Institutional Review Board (IRB) process of the University of Tennessee, prior to recruitment of participants. Prior to applying for the IRB review, the instructor of targeted course was approached via email and personally, to inform her of the proposed study, and its purpose, and to seek written approval for the conduct of study. Various logistical issues of scheduling and implementation, including access to Canvas – the course management software of the course for collecting data on student responses and scores on relevant quizzes, and tests, and for reviewing the course material related to studio 4, were decided through discussion and consensus, with the instructor and TAs. All relevant documents such as ‘informed consent’, ‘study information’, and ‘data collection instruments’ proposed to be administered to students, and information about the procedures and research plan were provided to the instructor, prior to beginning the IRB process.

After receiving IRB approval, I addressed the students verbally, during a 10-minute break of Studio Session 1 (Week 1), to inform them of the study, answer their queries, and to request their participation. I handed out copies of Information of the study, and consent form, and instructed them to sign the form and drop it in a closed drop box, kept on a table alongside the exit door, if they wished to participate. The ‘study information’ explained to the students that they would earn extra credit for participation, to be assigned to them at the end of semester, after all other grades had been assigned, based on a rubric, printed on the information document (Appendix 2A). The non-participants did not suffer, because other ‘extra credit’ opportunities in the form of ‘additional problems or activities’ were available to all students as standard practice in this course. By obtaining informed consent, providing a low stress environment, protecting data using pseudonyms, observing standard security protocols for file storage on computer, participants were treated ethically throughout the study, in accordance with all IRB guidelines of the university. Prototypes of ‘study information’ and ‘informed consent’ are appended in Appendices 2A, and 2B respectively.

After receiving informed consent, participants were instructed via email, to set up Twitter accounts with their institute assigned usernames and emails. During Studio Session 2, the setting up of accounts and a test run was conducted, to ‘walk’ students through the process of tweeting their thoughts, responding to specific prompts and reflecting on activities using the direct message mode. Twitter data was also collected during Studio 3, but not used in this study. The data collection for this study was implemented in four phases as described below (Figure 3.3 above). However, the same sample was examined in a temporal sequence in a single study, in four different inquiry contexts, using three different types of analytical techniques – quantitative, qualitative, and mixed-methods, with the purpose of consolidating the findings of each phase, to validate the proposed MM assessment mechanism.

Data Collection

In Phase 1, a validated, quantitative survey instrument – EBAPS version 5, assessing students’ epistemological stances with respect to physics and physics learning was administered using a web-based application, SurveyMonkey. The survey instrument is appended in Appendix 3A. In Phase 2, qualitative data on learners’ reasoning and framing of one Pre-Question, and two Post-Questions related to the content of the inquiry activity, was collected in Studio Session 4, for all four experiments. In Phase 3, student responses and scores on A3 Homework Quiz (pre-lab), Lab Quiz, Labs, and Test 1 were collected via Canvas for A3 Quiz and Test 1, and by making copies of students’ Lab Quiz and Labs submitted to the TAs.

Quantitative Data Phase 1: EBAPS

Per the Implementation Guide (Elby, Frederiksen, Schwarz, White, 2001), the Epistemological Beliefs Assessment for Physical Sciences (EBAPS) “probes the epistemological stances of High school and Introductory College students in introductory physics, chemistry and physical science” (p.2). The survey has a total of 30 items split into three parts, administered over a duration of 15-22 minutes. Part 1 has 17 statements which students are asked to rate on a Likert Scale of A) strongly disagree, B) Somewhat disagree, C) Neutral, D) Somewhat agree, E) Strongly agree. Part 2 contains 6 multiple choice opinion statements, and Part 3 contains 7 multiple choice conversation scenarios.

The five dimensions of students’ views probed by EBAPS are: (i) structure of knowledge – whether students view physics knowledge as a coherent unified whole, or as weakly connected pieces of facts and formulas (Q. 2, 8, 10, 15, 17, 19, 20, 23, 24, 28); (ii) nature of knowing and learning physics – whether they rely on relating new knowledge to prior experiences, intuitions, knowledge, by reflecting on their thinking, or on merely absorbing new information (Q. 1, 7, 11, 12, 13, 18, 26, 30); (iii) real-life applicability – whether they believe that physics knowledge and ways of thinking apply to real life, or only to classroom and laboratory contexts (Q. 3, 14, 19, 27); (iv) evolving knowledge - extent to which students view knowledge as absolute and set in stone, or as relative and evidence-based and/ opinion-based (Q. 6, 28, 29); (v) source of ability to learn – whether they believe that being good in physics depends on natural ability or on learning and doing physics (Q. 5, 9, 16, 22, 25) (Elby, Frederiksen, Schwarz, White).

Rationale. There were other stringently validated instruments available for assessing beliefs about physics such as the commonly used Maryland Physics Expectations Survey (MPEX) developed by Redish, Saul, and Steinberg (1998) of the PER Group at University of Maryland and the Colorado Learning Attitudes about Science Survey (CLASS) developed by Perkins, Adams, Podolefsky, Finkelstein, Dubson, Lemaster, Reid, ...and Wieman (2004) as part of the PhET (Physics Education

Technology) project, by the PER group of University of Colorado, Boulder. But EBAPS was chosen because of its theoretical alignment with the ‘resources model’ of cognition which framed this study. The designs of MPEX and CLASS assume that learners’ epistemological beliefs about physics exist as stable, coherent structures, while the items of EBAPS are developed to probe both - beliefs as transient pieces (resources) and as stable stances.

Scoring. Each item was scored on a scale of 0 (least sophisticated) to 4 (most sophisticated), as per the scoring scheme, which accounts for question-by-question variations in whether, for instance, neutrality is more, or less sophisticated. The survey is typically administered either as pre-post, or multiple choice, or agree-disagree. For the purposes of this study, it was administered pre-inquiry activities to assess learners’ prior epistemological beliefs. The survey instrument, and scoring scheme are appended in Appendix 3A, and 3B respectively.

Analysis. Student response data was manually entered into the EBAPS spreadsheet template, which automatically calculated student scores by subscale, by item, and/ by total. Overall scores were analyzed with the help of SPSS to identify low (below 33.33th percentile), medium (between 33.33th and 66.66th percentile) and high (above 66.66th percentile) scores. The results of grouping, and individual scores were used for quantitative correlation and validation of the ENA assessment, and for randomly picking 1 response from high score group, and 2 responses from medium and low score groups each, per question, per experiment, for rating of code categories by two independent raters.

Validity and Reliability of EBAPS. Items of EBAPS were synthesized through an extensive literature review of other epistemological surveys (Elby, Frederiksen, Schwarz, White, 2001). The developed survey instrument was administered to pilot subjects, and two sets of revisions were carried out based on informal feedbacks. Then it was administered to about 100 community college students, who were also asked to provide their reasoning for responding as they did to each question. Their responses were coded to exclude non-epistemological content and revise the questions. The reliability of the EBAPS categories was calculated, and results were acceptable in four of five categories. EBAPS has been given to over 1000 students in high school and introductory university physics courses at multiple-institutions, and the results have been published in over eight peer-reviewed publications. Reliability of most beliefs-based survey instruments is tested by ensuring that there is a high correlation as measured statistically, between the subscale items. But the developers of EBAPS chose not to use this technique, on the principle that they assume context sensitivity of beliefs, and do not subscribe to the assumption that subscales of beliefs are stable and consistent (Elby, Frederiksen, Schwarz, White, 2001).

Quantitative Data Phase 3: Course Quizzes, Tests, Labs

In Phase 3 of the study, quantitative scores of participants on 4 existing assessment instruments were collected for correlational analyses and validation of ENA assessment.

A3 Homework Quiz. This was a pre-lab homework Quiz with a total score of 8 points, containing eight questions – 3 multiple choice concept and /problem questions, 4 problems, and 1 conceptual question involving a diagram. The content covered was ‘electric circuits’, and the same Quiz was administered to all sections of the course. Students were allowed more than 1 attempt. Their scores on the first attempt, were considered for this study, since the final scores did not have meaningful variation, from the perspective of grouping students into low, medium and high score categories, and the first attempt was also more likely to mirror or reflect the influence of their contextual epistemologies, on their response. Students took the Quiz online, through Canvas, and could review the readings, when the system marked their response as wrong. Scores were obtained directly from Canvas.

Lab Quiz. This quiz was administered during the studio session, prior to the inquiry activities. Each quiz was scored by the TAs, for 10 points, and was also constructed by the TAs. So, Sections 1 and 3, had a different set of questions covering the same content. Both sets included problems on resistance and capacitance. Students’ quiz responses were photo-copied, after the quizzes had been graded by the TAs.

Labs. These were scored for 100 points by the TAs, and the same set of questions were administered to all sections of the course. Each of the four experiments, had questions about the findings from the experiment, and asked students to describe-explain-justify-predict results. Students’ Labs were photo-copied, after they had been graded by the TAs. This data was also used for triangulation purposes, to cross-verify the interpretations drawn from learners’ Twitter responses. The post questions were similar in content to the questions on labs, in terms of the reasoning asked of the students.

Test 1. This was a one-hour long, mid-term test, administered to all sections of the course, and scored for 15 points. Students were required to attempt 15 out of 17 multiple choice, conceptual and problem solving-questions. Like A3Quiz, the Test was autoscored by the course management system.

The A3 Home-work Quiz and Test 1 questions are appended in Appendix 4A, and 4B, respectively. Labs are appended in Appendix 1A to 1D.

Rationale. Using multiple contexts for testing utility of an assessment instrument, and comparison of performance with other assessment instruments are two criteria often used for validating an assessment instrument (Adams, & Weiman, 2011). Therefore, correlational analyses of performance outcomes of existing assessments and ENA-based assessment outcomes of sophistication, were used to

establish discriminant validity of ENA assessment. Since the study's rationale was that existing assessments were inadequate to assess sophistication of thinking, it was expected that there would be no correlation, or weak correlation, between existing assessment scores, and measures of sophistication, computed through this study.

Qualitative Data Phase 2: Twitter

Qualitative observations (data) of learners' thinking were obtained by sending question prompts and receiving students' responses to these questions, via Twitter. Twitter is a freely available Web 2.0 application that allows learners to share their thoughts and other resources such as pictures of diagrams captured by the cell phone camera, over multiple platforms such as tablets, laptops and mobile phones. In most instructional applications of Twitter, learners reflect upon or tweet their thoughts on specific aspects of a learning experience, such as while watching an educational video or movie; or in large lecture contexts, to engage students and to obtain feedback on learners' understanding and difficulties (Tiernan, 2014), or in the context of collaborating in teams to reflect on the concepts (Rankin, 2012; Young, 2010). The magnitude of student response to instructor's queries or twitter prompts is reported to have exceeded expectations in comparison to the relatively few students who would contribute otherwise, and the quality of responses is reported to be like, but more prolific than verbal responses (Paul, & Iannitti, 2012).

Rationale for using Twitter. For the purposes of this study, the primary outcome expected out of the selected data collection technology was to obtain data on learners' thinking in-context. To generate credible and substantive data on thinking, learners must be willing, and enthusiastic about using the technology to express, share and/report their thoughts. In a previous IRB approved study in the same setting and context, I constructed a survey consisting of 10 Likert-scale based multiple choice questions and 8 free response questions, to collect data on student perceptions of studio lab activities, over the entire Summer semester of 2015. The free response data collected, indicates that students chose to convey their thoughts in phrases, and short form, even to descriptive questions. So, it seems that technologies which present short forms of communication, are more likely to draw this population into an easy flow of conversation, and engage in sustained discourse, as proposed in this study. The choice of digital technology to collect data for this study, was therefore more student-centered than technology-centered, because the participant is in fact the source of the data, and the data cannot be retrieved credibly without the participant's active, motivated participation.

While tools such as Google forms, free response surveys (e.g. SurveyMonkey), or dedicated surveys created using Qualtrics and Blackboard/Canvas, and other course management tools, as well as social media tools such as Facebook, Socrative, Kahoot, and backchannel student response systems such as Today's Meet, Wordle, and ChimeIn, can serve as mechanisms to collect free response data on student

thinking, Twitter seemed to be a natural choice, because the 140-character limit enforces short, focused responses, which convey maximum meaning with minimal use of words. The persistent presence afforded by Twitter, allows for informal, free-flowing, just in time social interactions, in contrast to the formal discussions on forums or chats, which are part of course management systems, since the latter media required more stringent log-in protocols, and cumbersome navigation, to access the appropriate page to post a message (Dunlap and Lowenthal, 2009). Besides, from the point of view of implementation, Twitter offered the following advantages - availability on multiple platforms; no cost; user-friendly with low learning curve; and ready-to-use application programming interfaces (API) with Spreadsheet, for ease of data transfer and further analyses.

Future potential as a criterion. In terms of future potential, for refinements and automation of the assessment, Twitter offers the following advantages as compared to other social media platforms such as Facebook, and Instagram. (i) Twitter data is easier to gather, sort and retrieve, because of its hashtag norm. (ii) Twitter API is more open and accessible, making it easier for developers to create tools to access, process and retrieve data, thus increasing the availability of a wide range of tools for researchers. (iii) Twitter already has a rich library of APIs for R and Python programming environments, which can be used by programming novices such as educators, to develop domain-specific, customized machine learning applications for text analysis of collected qualitative data. (iv) Machine learning methods, which can enable the development of customized text mining and analytical applications, and qualitative techniques such as Sentiment analysis, may work better with Twitter because of the 140-character length of Tweets, which results in a relatively homogenous corpus of datum. (v) Twitter may be a viable option for generating data in a form that aligns well with Time series analyses of the natural flow of learners' thoughts with time, because of the real-time nature of Twitter posts (Ahmed, 2015; Burghardt, 2015).

Developers have already developed several initiatives at the back-end of Twitter, for data analysis. DMI-TCAT (Twitter Capture and Analysis Toolset) is one such promising open source initiative of the Digital Methods Initiative (DMI), an Internet Studies research group of the University of Amsterdam, which uses Twitter APIs to capture data, for computing Tweet statistics, and conducting content analysis and network analysis of tweets (Borra, & Rieder, 2014). Other examples include DiscoverText (Shulman, 2011), Netlytic (Gruzd, 2016), and Chorus (Brooker, Barnett, and Cribbin, 2016). Twitter could also work well as a data collection mechanism in the future, to conduct correlational studies of EEG (Electroencephalograph) or fMRI (functional magnetic resonance imaging) images of brain activation patterns, and ENA-based associational patterns of learners' thinking, in the context of responding to specific prompts related to physics. Such dual perspectives are critical for synthesizing bridges between educational neuro-sciences, and cognitive and learning sciences. The future research potential identified here, aligns with my plans to follow up this exploratory study, with iterative design

refinement cycles, in diverse contexts, and populations in physics. This consideration has been another principal driving factor, in the choice of Twitter as a data collection instrument.

Twitter Prompts. The format of instructor's prompt plays a significant role in directing learner's response (Junco, Heiberger, & Loken, 2011). Since the goal, in this study, was to assess shifts in learners' connections, during inquiry experiments, questions had to be crafted such that learners' ideas, understandings, and beliefs about the concept inquired through each experiment, could be elicited, before and after the experiment. Therefore, I studied and thoroughly reviewed the readings assigned for each experiment, and the structure, tasks and questions incorporated within the 'Labs', which consisted of reports of each experiment students were expected to submit after the studio session, in depth. I also had previously conducted a study in the same setting and context, to collect data on learners' perceptions about studio activities, using field observations and a survey consisting of closed and open-ended questions. Through the previous study, I had obtained item analysis data of all course assessments, which allowed me to get a sense of 'student difficulties' and 'problem areas'.

So, I used the lens of my prior research experiences in the same setting, my teaching experiences, and research-based recommendations to craft one pre-question, and two post-questions for each experiment. The recommendations which informed the design of questions are as follows: (i) Probing and knowing students' initial thinking, prior conceptions, and/ misconceptions, is necessary for addressing problematic conceptions, and building upon useful preconceptions (Minstrell, Anderson, and Li, 2011). (ii) Students need to interpret and explain real world data, such as observations recorded during inquiry, rather than decontextualized problems (Osborne, Duschl, & Fairbrother, 2002). (iii) Students need to make inferences, construct explanations, and give justifications, based on the observed data (Duschl, Schweingruber, & Shouse, 2007; Sandoval, & Reiser, 2004; NRC 2001). (iv) Students need to make generalizations in multiple inquiry and/problem solving contexts having slightly different features, that require use of the same concepts, to enhance transfer of knowledge (NRC 2001).

The pre-question was thus, designed to elicit information about learners' prior understandings of concepts they would be inquiring, as per recommendation 1 above. Post questions asked learners to describe-explain-justify and/draw inferences, about physical relationships associated with the concepts inquired, based on their observational data of experiments, per recommendations 2 and 3 above. Two post questions were asked to have students use the concepts twice, per the recommendation 4. Specific concepts inquired in each experiment, and the question prompts, are presented in Table 3.1 below. An Expert review process, which consists of having experts in the domain review the items on an assessment, is typically used for validating items on an assessment (Trochim, 2006, Adams, and Wieman, 2010). The questions were therefore sent to the four Teaching Assistants, who taught the Labs, a week before the

studio session, and discussed with them for ensuring their relevancy and appropriateness to inquiry context.

Twitter Mode. Data collection with Twitter, can be conducted using either the ‘Direct Message’ mode, or the ‘Tweet Mode’. In the Direct message mode instructors can send and receive responses from and to individual learners privately, and in the Tweet mode, a Group Twitter account is used so that all users can log into the account, and tweet their thought, questions, doubts, and/responses, to enable a whole class discussion (by projecting the twitter interactions on to the projection screen), without speaking out.

Table 3.1 Pre-Post questions and concepts inquired, by Experiment

Experiment	Concept	Question Prompt
E1PreQ	Resistance of human body	What do you believe will happen if you touch the leads of a voltage source with wet hand? Why?
E1PostQ		What did you learn about the relationship between resistance of your body and its wetness, from this experiment?
E2PreQ	Change in capacitance of a parallel plate capacitor	What are the ways by which you could change the capacitance of a parallel plate condenser? Elaborate.
E2PostQ1		What are the ways by which you changed capacitance in this experiment?
E2PostQ2		What did you learn about the relationship between capacitance of a parallel plate condenser and (a) distance between its plates and (b) area of the plates.
E3PreQ1	Ohmic and Non-ohmic materials	What are ohmic materials. Explain
E3PreQ2		What are non-ohmic materials. Explain.
E3PostQ1		Use your observations to decide if the bulb is an ohmic or a non-ohmic resistor. Explain your reasoning.
E3PostQ2		Use your observations to decide if the LED is an ohmic or a non-ohmic resistor. Explain your reasoning.
E4PreQ	Behavior of current and voltage in RC circuits	Current in a circuit containing resistance R connected to a voltage source, reaches the value V/R almost instantaneously. How will the behavior of current change when a capacitor is introduced in the circuit to make it an RC circuit?
E4PostQ1		Describe what you found about the behavior of current in an RC circuit from this experiment?
E4PostQ2		Explain the similarity and difference you found in the behavior of voltage and current in an RC circuit.

Chen and Chen (2012) used the Direct Message mode of Twitter, to collect data on student perceptions. Students evaluated each of the three weekly class sessions of a Research Methodology class, after each session, by responding to the questions - “(1) What points were most interesting? (2) What points were most confusing? and (3) What did you want to tell/ask about this session?” (p. E49).

Participants' responses were found to be useful as formative assessment tools, for evaluating instruction, and learning (Chen, and Chen, 2012). Another study, in which Twitter was used in the Tweet mode, reported positive outcomes in terms of increased classroom engagement, as measured by the content and volume of Tweets received (Tiernan, 2014). In this study, bot modes were initially planned to be implemented, but initial practice runs indicated two difficulties- (i) TAs were not able to handle simultaneously, Twitter interactions of participants, and face-to-face interactions of non-participants, and (ii) Participants had to send responses to two different accounts, for the two modes, and found it confusing to keep switching. Therefore, for logistical reasons, only the Direct Message Mode was used during Studio Session 4, which formed the context of this study.

Administration of assessment. Before, and after each experiment, one Pre-Question, and Two Post-Questions were sent as a group direct message (DM) from my Twitter handle, @ddeshpan to the Twitter accounts of all participants. The application, 'Multi DM', of TweetGuru.net, (The application has recently closed its service.) was used. It allows users to log into their Twitter accounts, and send the same message to 12 receivers at a time. The process was repeated twice for each set of pre-and post-questions. Students were asked to respond using the direct message mode of Twitter, since the 140-character limit does not apply to direct messages. Students' twitter responses (Direct messages) were downloaded after each Experiment, as a Spreadsheet file, using the application, twExList, of docteur-tweety.com. The application allows a maximum of 200 messages to be exported at a time. Each spreadsheet file downloaded, contained columns for date, sender account, and content of message, as shown below, in Figure 3.4.

1	Docteur Tweety - TWEXLIST - www.twexlist.com		
2	151 last direct messages received by @ddeshpan1		
3			
4	Date (GMT+1)	Direct message	By
5	9/6/2017 0:01	No worries!	@vhockada1
6	9/5/2017 23:50	#24.2 The way we found $S_1 - S_5$. $S_1 - S_2$ equalled 0. S_3 equalled -2. S_4 was +2 and S_5 was -1. The units used were nanocoulombs/ E0	@kadams3211
7	9/5/2017 23:50	#24.2 we answered them in lab $s_1=0$, $s_2=0$, $s_3=-2$, $s_4=+2$, and $s_5=-1$ in nanocolumbus e0	@lgentry71
8	9/5/2017 23:48	End exploration 3	@vhockada1
9	9/5/2017 23:48	Flux is just the charge of the area enclosed divided by a constant	@vhockada1
10	9/5/2017 23:47	Am I supposed to be able to see surfaces 1-5?	@vhockada1
11	9/5/2017 23:47	sorry that was the answer for 24.1	@kadams3211
12	9/5/2017 23:47	The net electric flux would entail what goes in and what comes out. If those are equal the net flux is 0. The flux is of whatever the surface is made of.	@vhockada1
13	9/5/2017 23:47	sorry #24.1	@lgentry71
14	9/5/2017 23:47	#23.23 The factors on which an electric flux depend on the charge and number of charges	@kadams3211
15	9/5/2017 23:47	#22.23 it depends on the charges and particles	@lgentry71
16	9/5/2017 23:45	both questions did agree with my expectations \	@lgentry71
17	9/5/2017 23:45	# 23.22 with the multiple charges it created different direction of lines.as in making almost circular lines	@lgentry71
18	9/5/2017 23:43	#23.21 when I put the negative charges on the electric field all of the lines pointed in or straight into the negative charge. For positive the same thing happened except for all of the lines went out.	@lgentry71
19	9/5/2017 23:42	#23.22 The electric field was stronger at some points and weaker at others. The field lines were distributed unevenly. My observations agreed with my expectations	@kadams3211

Figure 3.4. Screenshot of learners' Twitter responses downloaded using TWEXLIST.

Qualitative Analysis Phase 2: Spreadsheet

Qualitative data analysis software applications (QDAS) have witnessed a transition from traditional stand-alone packages such as NVivo, MaxQDA, and Atlas.Ti, which developed in the 1980s, to QDAS 2.0, or web-based applications such as Dedoose, to Cloud QDA such as DiscoverText. While the traditional QDAS still outperform the web and cloud-based applications, in terms of the breadth and depth of qualitative analysis features they provide, the latter offer many benefits, such as low cost, low learning curve, ease of integration with quantitative tools and statistical analyses. and customization potential (Denzin, & Giardina, 2016). Thus, the latter applications would be ideal choices for qualitative analyses proposed in this study. But manual coding on Spreadsheet was chosen for qualitative coding in this study, mainly because it presented a good fit for further coding of co-occurrences of each code category as '0' or '1' for all responses, and further sorting, merging, and processing of the data. The qualitative data collected using Twitter, was exported to Spreadsheet for qualitative analysis. The basic purpose of qualitative analyses was to uncover and interpret how meaning was constructed by learners (Merriam, & Tisdell, 2015); interpreting how learners made sense of physics concepts.

Procedures. Responses for each question were first processed and coded on a separate worksheet, and then merged together, along with other metadata (e.g. scores on various assessments and corresponding group assignment, Section, Gender) onto one worksheet. Each row of Spreadsheet, thus represented one learner's (case) response to either a Pre-or post-question. The process of coding each response is elaborated separately in the following paragraphs. After all data was coded, I examined the code categories by question, and by Experiment, to identify dominant patterns of knowledge elements in learners' thinking, to answer RQ1. Prior research, on concept formation and epistemic framing informed the identification of themes for this study. Concepts in physics require the use of verbal, symbolic, mathematical and schematic representations to convey or reveal a detailed picture of underlying structural and dynamic relationships which govern physical phenomena, to learners. In turn, learners are required to process multiple representations of knowledge in parallel, while also establishing connections to prior knowledge, for internalizing and constructing new knowledge. Since different learners tend to process the same information using different forms of representation in physics, the categorization of knowledge elements, from general to specific codes (Burns & Grove 2005), was largely based on the knowledge representations chosen by learners.

Content Analysis. The specific method of analysis used for qualitative analysis, was directed content analysis. Content analysis has been variously described as – a technique to make replicable, valid inferences from texts to the contexts of their use (Krippendorff, 2004); a technique for objective, systematic and quantitative description of the content of communication (Berelson, 1952); a systematic, objective means to make valid inferences from verbal, visual, or written data to describe and quantify

phenomena (Downe-Wambolt, 1992). While it is generally, classified as qualitative and quantitative (Bengtsson, 2016), Hsieh and Shannon (2005) classify content analysis as – conventional (inductive), directed (deductive), and summative (quantitative), based on the approach to coding. Prior to discussing the coding technique, I will elaborate upon the unit of analysis which guided the analysis process.

Unit of analysis. The choice of sampling unit and unit of analysis should be informed by the research's focus; there are no established criteria in content analysis to determine the size of a sampling unit, or unit of analysis (Bengtsson, 2016; Krippendorff, 2004; Patton, 2005). The sampling unit for this study consisted of individual learners, and the unit of analysis for qualitative analyses, was the content of an individual learners' response to a specific question. The response contained one or more lines, and/ if more than one response was sent by a participant, the combined response was analyzed.

Coding. Coding is the process of assigning codes to the data, where “A code is most often a word or short phrase that symbolically assigns a summative, salient, essence-capturing, and/or evocative attribute for a portion of language-based or visual data.” (Saldana, 2009, p.2) In directed or deductive content analysis, existing theory or prior research, is first used to identify key concepts or variables as initial coding categories (Potter & Levine-Donnerstein, 1999; Hsieh & Shannon, 2005), and operational definitions are determined based on theory. Since the size of unit of analysis in this study was small, coding began immediately on reading each response, with initial codes driven by learners' own words, which were often similar across participants, since all participants were responding to the same concepts, and were interpreting the same readings. I read each response, mentally identified the knowledge elements chosen by the student, and coded it almost verbatim, retaining the main content. So, code columns were added to the spreadsheet containing the raw data, as each new response to a question, was analyzed, if needed. Initial coding resulted in about twenty to thirty units of code per question.

A cell (i, j) was assigned the value 1, if the code assigned to column j, was present in the content of response corresponding to record (row) i. Once all 1's had been assigned, the remaining cells were filled with 0's, by selecting the area of spreadsheet to be assigned numerical values, then clicking on the 'Find & Special' tab of spreadsheet at top right, selecting the 'Go to Special' option, choosing 'blanks', filling in the value '0' in the highlighted cell, and pressing 'control + enter'. After all the 38 responses, for each question, were coded, similar code categories were merged together through an iterative process (Burns & Grove 2005). Frequency counts of each code category were calculated using the 'auto sum' feature of Spreadsheet, and if the total count was found to be less than 2, the code columns were either eliminated, or merged with 'similar codes'.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V					
	E.Change K.	Change se	K.Change K.on	swit	K.Change K.	Change K.	Change K.fully	char	K.Change K.	Change K.	chang K.	IncV	K.	Change K.	add R.	K.T=RC	K.change s	K.change sh	K.llel	Zser	K.No	Plat	K.Orient	K.series	or	l
To change capacitance, you could change plate separation and charge capacitance of the parallel plate. also the plate spacing of the capacitors and the area of the capacitors.	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
In order to change the capacitance of a parallel plate condenser, I guess you could change amount of resistance in the plate.	1	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
To change capacitance, you could change the voltage or change the distance between the plates	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Capacitance can be changed by increasing or decreasing voltage, or resistance	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
To change capacitance you can change current or voltage.	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
To change capacitance change the distance of them. Or increases the voltage?	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
you could change it by the different amounts of volts, weight, and distance you are measuring?	1	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
you can change the capacitance by adding more resistors. I think this is right because I know that T= R (resistance) * C (Capacitance).	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	
To change capacitance you can change the size and shape of the plates. You could also change the distance between the plates.	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	
You can change capacitance by changing the capacitor to a series plate instead of a parallel	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	
you can change the size or shape of the circuit. The distance between the plates could also change the capacitance.	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	
You can change the size or shape of the circuit. The distance between the plates can also change the capacitance. The number of plates and how they are oriented to one another in space can change too.	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	1	1	0	0	0	0	
To change capacitance, you could change whether the plate is series or parallel	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	

Figure 3.5. Example responses and 21 code columns for PreQ of Experiment 2.

Figures 3.5 to 3.8, illustrate the process of consolidation (merging). Figure 3.5 above, presents a screen shot of 14 responses to PreQ of Experiment 2, which asked “What are the ways by which you could change the capacitance of a parallel plate condenser?”

The portion of spreadsheet shown above is a partial view of the entire worksheet, and contains 21 out of a total of 30 initial code columns. Code column C (K. Change separation) in figure 3.5 above, was merged with column I (K. Change distance), because ‘separation between plates’ can be treated as same in meaning to ‘distance between plates’, in physical terms. Merging was such, that if one or all merged columns had a cell value of 1, then resultant column had a value of 1 for a given record or row.

A	B	C	D	E	F	G	H	I	J	K	L
	E.Change C.	K. Change d	K.Change Q	K. Change A	K. Change R	K. Change V	K. change I	K. Change W	K.T=RC	K.change si/sh	K.changeN/Orient
To change capacitance, you could change plate separation and charge	1	1	1	0	0	0	0	0	0	0	0
Capacitance can be changed by flipping a switch on, fully charging the capacitance of the parallel plate. also the plate spacing of the capacitors and the area of the capacitors.	1	1	1	1	0	0	0	0	0	0	0
In order to change the capacitance of a parallel plate condenser, I guess you could change amount of resistance in the plate.	1	0	0	0	1	0	0	0	0	0	0
To change capacitance, you could change the voltage or change the distance between the plates	1	1	0	0	0	1	0	0	0	0	0
Capacitance can be changed by increasing or decreasing voltage, or resistance	1	0	0	0	1	1	0	0	0	0	0
To change capacitance you can change current or voltage.	1	0	0	0	0	0	0	0	0	0	0
To change capacitance change the distance of them. Or increases the voltage?	1	1	0	0	0	1	0	0	0	0	0
you could change it by the different amounts of volts, weight, and distance you are measuring?	1	1	0	0	0	1	0	1	0	0	0
you can change the capacitance by adding more resistors. I think this is right because I know that T= R (resistance) * C (Capacitance).	1	0	0	0	1	0	0	0	1	0	0
To change capacitance you can change the size and shape of the plates. You could also change the distance between the plates.	1	0	0	0	0	0	0	0	0	1	0
You can change capacitance by changing the capacitor to a series plate instead of a parallel	1	0	0	0	0	0	0	0	0	0	1
you can change the size or shape of the circuit. The distance between the plates could also change the capacitance.	1	1	0	0	0	0	0	0	0	1	0
You can change the size or shape of the circuit. The distance between the plates can also change the capacitance. The number of plates and how they are oriented to one another in space can change too.	1	1	0	0	0	0	0	0	0	1	1
To change capacitance, you could change whether the plate is series or parallel	1	0	0	0	0	0	0	0	0	0	0
	14	7	2	1	3	4	0	1	1	3	2

Figure 3.6. Result of consolidation of 21 columns of Figure 3.5.

Columns corresponding to codes ‘K.fully charged’ and ‘K. on switch’, were merged with ‘K. charge Q’, since the student uses the terms ‘flipping the switch on’, together with ‘fully charging’, and the

content can therefore be interpreted as same in meaning to ‘increase and/ decrease of charge’ or change in charge (Q) of the capacitor. Similarly, codes ‘K.Change Size’ and ‘K.Change shape’ were merged together, since they approximately referred to the same physical dimensions, for a parallel plate capacitor; codes ‘K.add resistance’ and ‘K.change resistance’ were merged because they meant the same thing physically; ‘K.Increase Voltage’ and ‘K.Change Voltage’ were merged together under ‘K.Change Voltage’, because ‘increase in voltage’ was subsumed by ‘change in voltage’. Resulting 11 columns are shown above in Figure 3.6.

Figure 3.7 shows the results of further consolidation, in which K.Change Q/V/I/V subsumed four code columns, because physically speaking, change in either charge, voltage, current, and/ resistance, will result in change in the other quantities. Code ‘K.Change Circuit/PlateSize/Shape/Orientation/Weight’, refers to change in physical dimensions, and subsumes change in weight which will result when either size, shape, number of plates, and any other structural changes are carried out.

A	B	C	D	E	F	G
	E.Change C	K.Change Q/I/V/R	K. Change A	K. change d	K. Change Cct/PI/Si/Sh/Or/W	K.T=RC
To change capacitance, you could change plate separation and charge	1	1	0	1		0
Capacitance can be changed by flipping a switch on, fully charging the capacitance of the parallel plate. also the plate spacing of the capacitors and the area of the capacitors.	1	1	1	1		0
In order to change the capacitance of a parallel plate condenser, I guess you could change amount of resistance in the plate.	1	1	0	0		0
To change capacitance, you could change the voltage or change the distance between the plates	1	1	0	1		0
Capacitance can be changed by increasing or decreasing voltage, or resistance	1	1	0	0		0
To change capacitance you can change current or voltage.	1	0	0	0		0
To change capacitance change the distance of them. Or increases the voltage?	1	0	0	1		0
you could change it by the different amounts of volts, weight, and distance you are measuring?	1	0	0	1		1
you can change the capacitance by adding more resistors. I think this is right because I know that $T = R \text{ (resistance)} * C \text{ (Capacitance)}$.	1	1	0	0		1
To change capacitance you can change the size and shape of the plates. You could also change the distance between the plates.	1	0	0	1		1
You can change capacitance by changing the capacitor to a series plate instead of a parallel	1	0	0	0		1
you can change the size or shape of the circuit. The distance between the plates could also change the capacitance.	1	0	0	1		1
You can change the size or shape of the circuit. The distance between the plates can also change the capacitance. The number of plates and how they are oriented to one another in space can change too.	1	0	0	1		1
To change capacitance, you could change whether the plate is series or paralel	1	0	0	0		1
	14	6	1	8		6

Figure 3.7. Result of consolidation of columns in Figure 3.6.

Thus, 11 code columns were merged, resulting in six columns (Figure 3.7) for the partial worksheet illustrated. This process was replicated for all 12 questions. Then the final codes were examined, for similarities and differences, and it became clear, that for each set of codes corresponding to one worksheet, four broad categories, or themes could be identified, based on the type of knowledge, and the source of knowledge.

	A	B	C	D	E	F	G
1		E.Change	K3.Change	K1. Change	K1.change	K3. Change	K2.T=RC
2	To change capacitance, you could change plate separation and charge	1	1	0	1	0	0
3	Capacitance can be changed by flipping a switch on, fully charging the capacitance of the parallel plate. also the plate spacing of the capacitors and the area of the capacitors.	1	1	1	1	0	0
4	In order to change the capacitance of a parallel plate condenser, I guess you could change amount of resistance in the plate.	1	1	0	0	0	0
5	To change capacitance, you could change the voltage or change the distance between the plates	1	1	0	1	0	0
6	Capacitance can be changed by increasing or decreasing voltage, or resistance	1	1	0	0	0	0
7	To change capacitance you can change current or voltage.	1	0	0	0	0	0
8	To change capacitance change the distance of them. Or increases the voltage?	1	0	0	1	0	0
9	you could change it by the different amounts of volts, weight, and distance you are measuring?	1	0	0	1	1	0
10	you can change the capacitance by adding more resistors. I think this is right because I know that $T = R \text{ (resistance)} * C \text{ (Capacitance)}$.	1	1	0	0	0	1
11	To change capacitance you can change the size and shape of the plates. You could also change the distance between the plates.	1	0	0	1	1	0
12	You can change capacitance by changing the capacitor to a series plate instead of a parallel	1	0	0	0	1	0
13	you can change the size or shape of the circuit. The distance between the plates could also change the capacitance.	1	0	0	1	1	0
14	You can change the size or shape of the circuit. The distance between the plates can also change the capacitance. The number of plates and how they are oriented to one another in space can change too.	1	0	0	1	1	0
15	To change capacitance, you could change whether the plate is series or paralel	1	0	0	0	1	0
16		14	6	1	8	6	1

Figure 3.8. Resulting theme categories obtained after consolidation and categorization of 21 columns of Figure 3.5.

The thematization result of code columns of Figure 3.7, is shown in Figure 3.8. The process of finding ‘themes and categories’ will be elaborated in Chapter 4, which presents the Findings and Results if all analyses. The next phase of analysis (phase 3), network analysis, was conducted using ENA Web tool 4.0, which requires that data be inputted in a specific format, as a .csv file. Hence columns of all 12 spreadsheets, were organized as ‘metadata’ (Figure 3.9 below), ‘raw data’ (Figure 3.10 below), and ‘code columns’ (Figure 3.11 below). The following 3 figures illustrate how the spreadsheet looked like, for each experiment and question. Columns A to AD contain metadata (Figure 3.9), column AE contains the raw data (Figure 3.10), and columns AF to AT contain the consolidated code categories for Experiment 1, PreQ and PostQ (Figure 3.11).

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	
Date	Student	Gender	Sector	Studio	Exper	Questic	axis_1	axis_2	axis_3	axis_4	axis_5	overall	EBAP56	A3Attent	A3(J)Grc	A3Quiz	A3Group	Quiz4	Lab4Grc	H3Attent	H3(J)Grc	H3Quiz	H3Group	Test1	Test1Grc	Group	Username			
19-Sep F23	F	1	4	1	PreIQ	1.75	2	1.75	2.333333	0.8	1.766667	High	7	Medium	7	Medium	8	High	91	Low	8	High	7	Medium	2	Low	F2	@Kaylaf		
19-Sep F26	F	1	4	1	PreIQ	1.45	1.375	1	1.666667	1.8	1.636667	Medium	7	Medium	8	High	7	Medium	8	High	99	High	8	High	5	Low	E	@kweim		
19-Sep F27	F	3	4	1	PreIQ	1.25	1.875	1.75	1.333333	2	1.636667	Medium	7	Medium	8	High	7	Medium	7.5	Medium	101	High	8	High	6	Low	K2	@lgentry		
19-Sep F38	F	1	4	1	PreIQ	0.8	1.4375	1.375	1	2.8	1.466667	Low	7	Medium	8	High	8	High	6	Low	82	Low	7	Medium	7	Low	A	@Samar		
19-Sep F1	F	1	4	1	PreIQ	1.7	1.375	1	2.666667	14	1.766667	High	8	High	8	High	8	Medium	84	Low	8	High	8	High	7	Low	S	@lastua		
19-Sep F21	F	3	4	1	PreIQ	1.75	1.625	1.5	1.666667	14	1.686667	Medium	6	Medium	8	High	10	High	36	Medium	8	High	8	High	8	Low	L	@jletzler		
19-Sep F41	F	3	4	1	PreIQ	1.5	1.25	2	1.333333	2	1.5	Low	8	High	8	High	10	High	98	High	8	High	8	High	8	Low	J	@sseibe		
19-Sep F16	F	1	4	1	PreIQ	1.7	1.5625	1.75	1	1	1.586667	Medium	6	Medium	8	High	3	Low	76	Low	8	High	8	High	9	Low	X1	@gpitre		
19-Sep M2	M	1	4	1	PreIQ	1.1	2.375	1.5	1	2	1.686667	Medium	8	High	8	High	3	Low	81	Low	8	High	8	High	9	Low	F2	@cyoho		
19-Sep F5	F	1	4	1	PreIQ	1.85	0.875	2	2.333333	24	1.75	High	6	Medium	6	Low	4	Low	82	Low	8	High	8	High	9	Low	A	@ahom		
19-Sep F20	F	1	4	1	PreIQ	1.9	1.75	2.25	1	1.6	1.9	High	8	High	8	High	8	High	9	High	83	Low	8	High	8	High	9	Low	P	@i_cook
19-Sep F13	F	1	4	1	PreIQ	1.95	1.0625	1.875	0.666667	3	1.632333	Medium	7	Medium	8	High	8	High	8	High	88	Medium	8	High	8	High	9	Low	H	@eevan
19-Sep F14	F	1	4	1	PreIQ	1.5	1.125	1.625	1.666667	2.6	1.65	Medium	8	High	8	High	8	Medium	88	Medium	8	High	8	High	10	Medium	H	@ehamr		
19-Sep F32	F	3	4	1	PreIQ	1.6	1.375	2	1.666667	24	1.7	High	5	Low	7	Medium	8.5	Medium	92	Medium	8	High	8	High	10	Medium	D	@oorozc		
19-Sep F33	F	1	4	1	PreIQ	14	2	1.25	1.333333	3.2	1.766667	High	7	Medium	7	Medium	10	High	93	Medium	8	High	8	High	10	Medium	G	@pelwv		
19-Sep F36	F	1	4	1	PreIQ	1.45	1.375	1.875	1.666667	2.8	1.7	High	7	Medium	8	High	10	High	96	Medium	8	High	8	High	10	Medium	C	@rwatso		
19-Sep F28	F	3	4	1	PreIQ	1.1	1.75	1	1.666667	2	1.466667	Low	8	High	8	High	7	Low	100	High	8	High	8	High	10	Medium	W	@lwidem		
19-Sep F9	F	3	4	1	PreIQ	0.777778	1.74286	1.75	2	2.25	1.702104	High	8	High	8	High	10	High	100	High	8	High	8	High	10	Medium	W	@cmabe		
19-Sep F4	F	1	4	1	PreIQ	1.1	1.375	1	2	1.2	1.233333	Low	7	Medium	8	High	4	Low	84	Low	8	High	8	High	11	Medium	S	@aridge		
19-Sep F2	F	1	4	1	PreIQ	1.7	1.875	1.75	1	24	1.766667	High	7	Medium	8	High	8	Medium	84	Low	8	High	8	High	11	Medium	B	@fowor		
19-Sep F10	F	1	4	1	PreIQ	1.95	1.5	2.125	1.333333	14	1.632333	Medium	8	High	8	High	8	Medium	93	Medium	8	High	8	High	11	Medium	G	@coubb		
19-Sep F35	F	3	4	1	PreIQ	1.85	1.25	2.5	2.333333	0.8	1.583333	Medium	8	High	8	High	10	High	96	Medium	6	Medium	6	Low	11	Medium	L	@ramst		
19-Sep F12	F	1	4	1	PreIQ	1.5	1.4375	1.875	2	24	1.686667	Medium	8	High	8	High	8	High	6.5	Low	98	High	8	High	11	Medium	E	@eellis2		
19-Sep F40	F	3	4	1	PreIQ	2.25	1.625	1.75	2	2	1.88	High	3	Low	5	Low	7	Low	100	High	1	Low	8	High	11	Medium	W	@smrii		
19-Sep F37	F	3	4	1	PreIQ	1.5	1.875	1.25	1	3	1.65	Medium	8	High	8	High	7	Low	103	High	8	High	8	High	11	Medium	F2	@rwood		
19-Sep F15	F	1	4	1	PreIQ	1.6	1.5	2.5	1.333333	12	1.586667	Medium	8	High	8	High	8	High	7	Low	82	Low	7	Medium	8	High	12	Medium	A	@gnanc
19-Sep F11	F	1	4	1	PreIQ	1.5	1.375	1	1.333333	24	1.533333	Medium	2	Low	5	Low	8	Medium	92	Medium	2	Low	7	Medium	12	Medium	D1	@dwash		

Figure 3.9. Metadata columns (A to AD) containing student scores on different instruments, and the corresponding group membership.

AD	AE
Username	
@Kaylaf	#pre1If you touch a voltage with wet hands it makes your hands more likely to conduct a charge and as a result you're going to get shocked
@kweim	If you touch a voltage source with wet hands you will most likely be electrocuted because water is a conductor of electricity.
@lgentry	#pre1if you touched a voltage source with wet hands you would get electrocuted, I am not sure as to why but that is what would happen
@Samar	#pre1You will get shocked. Water increases the available number of electrons. The charge will jump and shock you when you were wet.
@lastua	#pre1you would get shocked because of the electrical course
@jletzler	#pre1 The voltage would be transferred to your hand resulting in you being shocked
@sseibe	#pre1The voltage will vary most likely be lower than with dry hands
@gpitre	#pre you can charge it by using knowing the voltage
@cyoho	Distance lowers he capacitance and witness lowers the current of the thumbs
@athom	#pre1 if you touch a voltage source with wet hands the charge will transfer some energy to you because you are able to more readily conduct that energy from @athom91
@i_cook	#pre1 you will get shocked because wet hands provide a conducting path to the current to flow inside the body from the wire
@eevan	The voltage would be greater because it will be able to flow through easier
@ehamr	The voltage would increase because water easily carries the electrical current
@00FOZC	#questionpre1 you will experience some form of electric current with wet hands. Your body's resistance is lower so you will experience more of a shock with wet hands than dry

Figure 3.10. Raw data column (AE) containing the text of student's Twitter response to one question of one experiment.

AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT
.....	K2.Shoc	K2.VIIIQ	K3.VIIIQ	K2.WCo	K2.WNC	K1.wetne	K1.wHBette	K1.Wetn	K1.Wetn	K1.dryH	K1.I.Pvar	K2.Nosh	K3.Wetr	E.Interpret_Rel/Wetnes
1	1	0	0	0	0	0	0	1	0	0	0	0	0	0
1	1	0	0	0	1	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	1	0	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	1	0	0	0	0	0	0
1	1	0	0	0	0	0	0	1	0	0	0	0	0	0
1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	1	0	0	0	0	1	0	0	0	0	0	0
1	1	0	0	0	0	0	0	1	0	0	0	0	0	0
1	0	1	0	0	0	0	0	1	0	0	0	0	0	0
1	0	1	0	0	0	0	0	1	0	0	0	0	0	0
1	0	1	0	0	1	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	1	0	0	0	0	0	0	0

Figure 3.11. Code columns (AF to AT) containing the code for each code category (0 = absent, 1 = present).

Network Analyses Phase 3: ENA Web Tool

Data in ENA format has four variables in separate columns. (i) Unit Variables are the units of analysis for which ENA will build network models e.g. networks from different learners, or from people with different conditions (EBAPS scores, gender, and other data columns with scores) (ii) Conversation variables can be units of time (activity separated into time segments), steps in a process (e.g. activity), or any way of identifying a unit in the data for quantifying relations between units. Units in lines in the same conversation are related to one another in the model. (iii) Code variables, which have binary values, are all the epistemic frame elements identified during coding. (iv) Raw data variables will include coded raw data excerpts, person who tweeted, time when the person tweeted. They are automatically included in the ENA set created, and do not have to be selected.

Procedures

The ENA Web Tool was accessed at <http://www.epistemicnetwork.org/live>. To access it, an account was created by signing up, with a username and password. All data files were uploaded to the application – one consolidated file for each experiment, for Pre-post network comparisons on one graph; and a separate file for each question, for detailed and centered network graphs of each experiment, and to generate a separate rData file for each experiment, for further quantitative analyses. The basic steps to launch the application, create ENA sets, and plot ENA sets are outlined below (Figures 3.12 to 3.15).

Create ENA sets. (i) Login with username and password (left screen Figure 3.12). (ii) Click on the upload ‘cloud’ symbol in the right screen (iii) Select the .csv file to be uploaded from computer and click upload. (the right screen shows uploaded files). (iv) Click on the file whose data needs to be modeled as ENA set. Figure 3.13 shows the screenshot of how the screen looks like when the file Expt1PreQ.csv was selected. (v) Next, ‘student ID’ was selected for the units and conversations, and code units were selected. The resulting screen can be found in Figure 3.14, which shows the name E1Pre filled in for the ENA set to be created, and the ‘create ENA set’ button activated. (vi) On clicking the create ENA set button, the ENA Tool opens a new screen (Figure 3.15), which shows all the ENA sets created (previously).

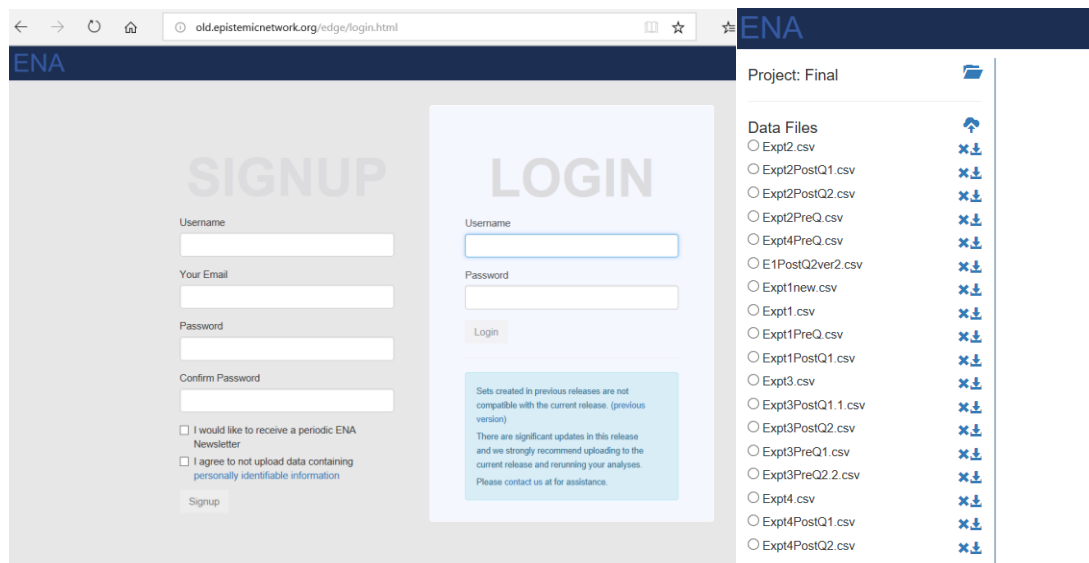


Figure 3.12. ENA Login Screen.

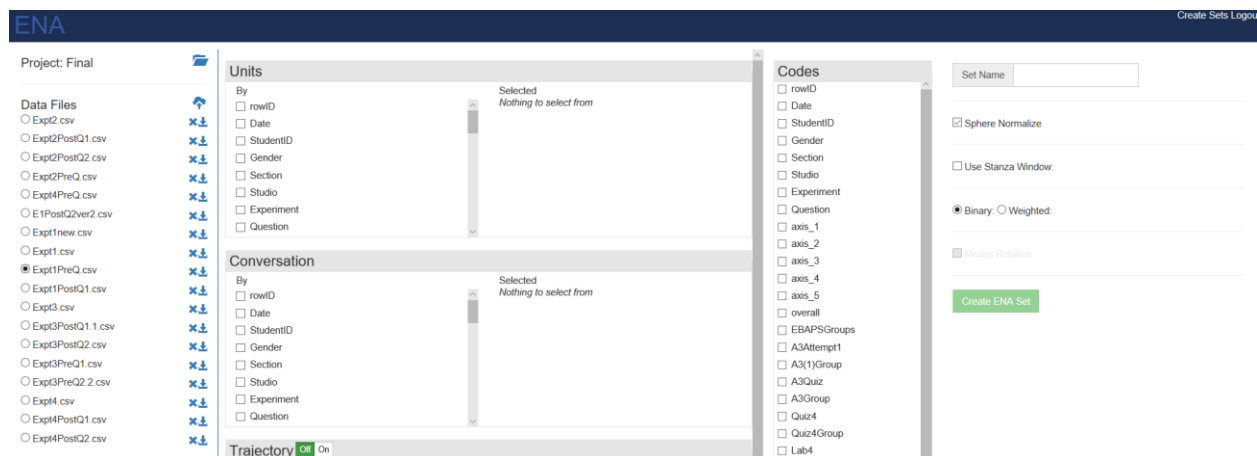


Figure 3.13 . The units and conversation fields get populated with the metadata, and the codes field gets populated with codes and some data columns, when a dataset (Expt1PreQ.csv) is chosen from the data files on left, as shown.

Plot ENA sets. (i) Select ENA set to be plotted. (ii) Click “Plot” and a new pop-up window will open (Figure 3.16 without the red centroids). (iii) Add items from the “Item Bank” to “Priority List” by dragging them, for more grouping options. (iv) Drag and reorder items in “Priority List” to change the selection order under “Selection Tree.” (v) Select desired units and their colors, and check Data1 at the top of screen to plot the centroids (locations of individual student’s epistemic network for PreQ of Experiment 1, as per the current choice of units (Figure 3.16). (vi) Click on a centroid, to view the epistemic network connections of a student. (vii) Click on the mean centroid (the solid square at center of graph in Figure 3.16) to view the mean epistemic network of all students (Figure 3.17).

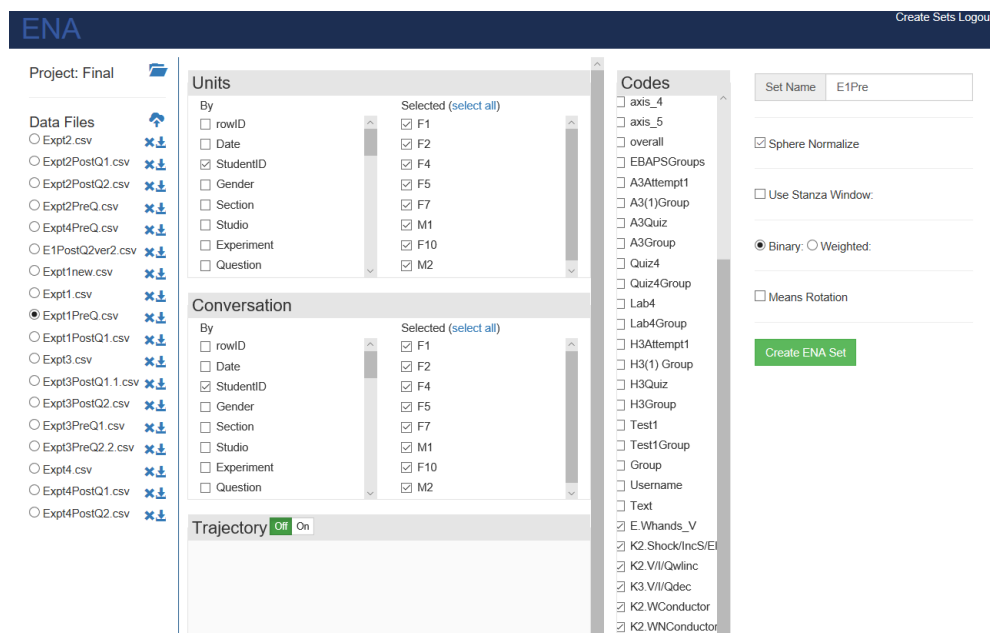


Figure 3.14. Screen with ‘Create’ button activated, when the units, conversation, codes are selected, and the ENA set is given a name.

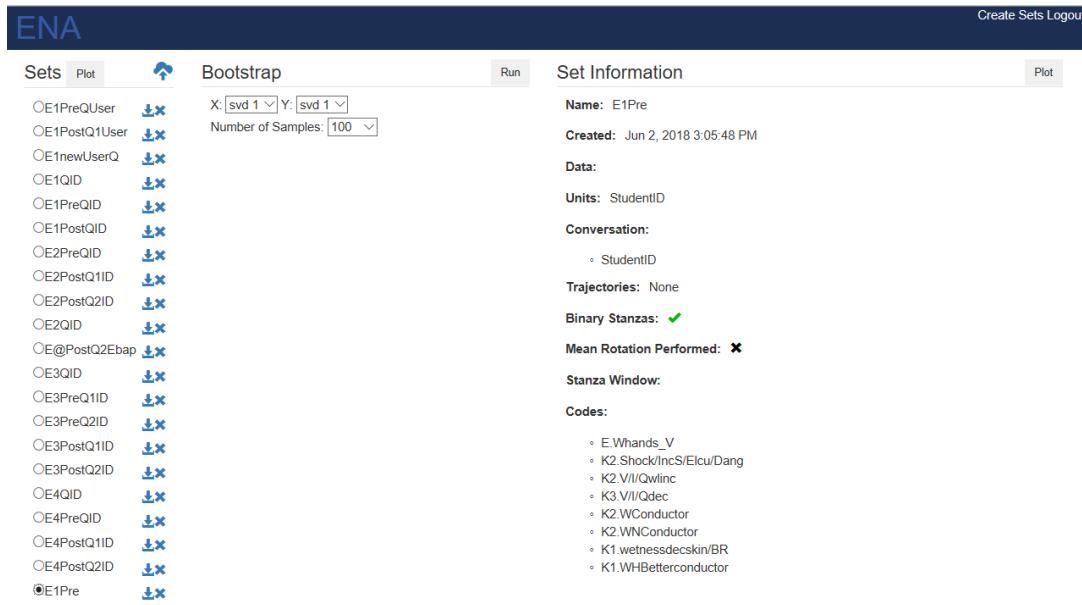


Figure 3.15. Plot screen which shows the ENA sets created, and details of entities – units, conversation, codes, of the ENA set chosen to be plotted (E1Pre).

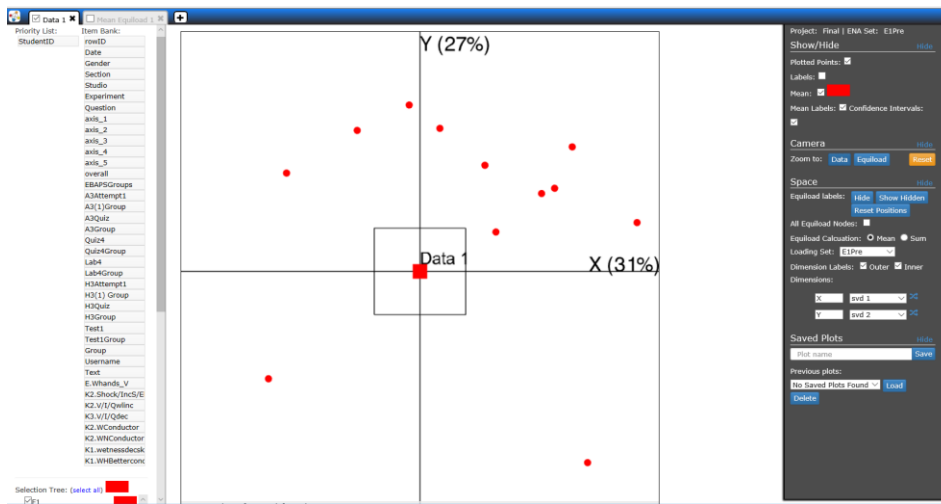


Figure 3.16. The network space which shows the plot of centroids (mean network of one student ID) of all units selected in the selection tree.

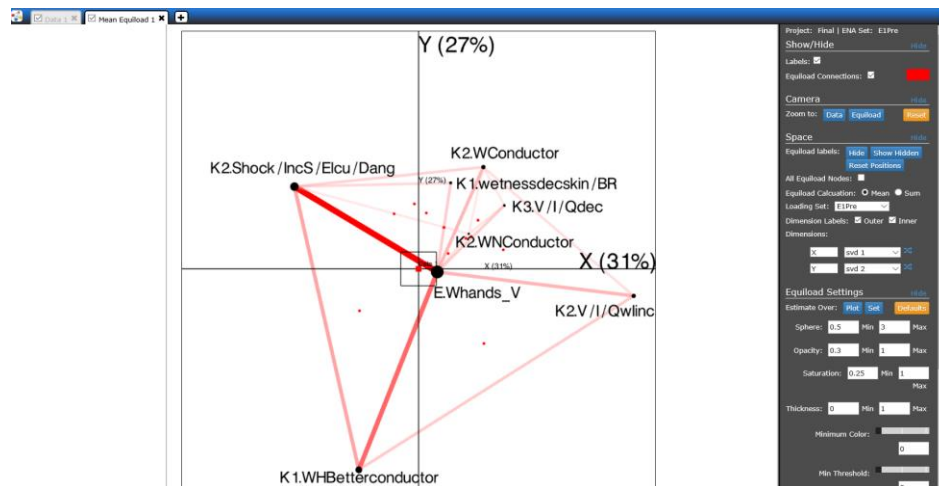


Figure 3.17. Plot of the mean Epistemic Network of all units selected.

Analyses. The following analyses were conducted for each experiment and each question:

(1) The consolidated file for each experiment, was used to plot and compare mean networks of students before experiment (PreQ), and after the experiment (PostQ1 and PostQ2), to answer RQ 2(a). The ENA web Tool allows testing for statistically significant differences using t-test.

(a) T-tests were conducted for Pre-post networks, along with qualitative observations of the strength of connections (based on the thickness, length, and/ brightness of the edges or lines) and types of node (the themes or code categories that were found to occur or were missing from learners' network connections).

(b) Network graphs showing centroids, confidence intervals of means, (like Figure 3.16), and mean networks (like Figure 3.17) for Pre-and Post-questions in different colors, were cross-checked to obtain qualitative information about differences in Pre-Post networks. For instance, by cross-checking Figures 3.16, and 3.17, it can be broadly determined that students whose centroids are found in quadrant (+X+Y) of graph in Figure 3.16, made most connections to the nodes (knowledge elements) found in quadrant (+X+Y) of graph in Figure 3.17.

(c) Pre-post network trajectories of individual students were examined, and the qualitative interpretations were drawn upon, to interpret the shift in their networks.

(2) The individual file for each question, was plotted and analyzed to determine differences in mean networks of students (a) by section, and (b) by low, medium high score groups for (i) EBAPS (ii)

A3HomeworkQuiz (iii) Lab Quiz (iv) Labs, and (v) Test 1 scores. For each assessment instrument, T-tests were conducted and qualitative examinations of mean networks of low, medium, and high score groups were conducted, to answer RQ 2(b), and to establish concurrent validity of ENA assessment by confirming that the ENA assessment could discriminate between mean networks of students grouped by their scores on five existing assessments used in the course.

Quantitative Analyses Phase 4: RStudio

Since one of the goals of this study was to explore quantitative metrics which could be utilized for formative assessments of sophistication of thinking, in introductory university physics courses, the rData file of each ENA set, created as seen on the left side of Figure 3.15 above, was downloaded on to the computer, and then imported into the RStudio environment for further processing, before computation of sophistication scores. RStudio is an open source statistical computing software package, with several a base programming language and standard commands. Since 12 rData files had to be processed, using repetitive complex commands, I created an rScript (Appendix 9), which could be run quickly for each file, to obtain output in the form of csv file, which was exported to the computer. The csv outputs for each question (of each experiment) were then saved as Excel files and the sophistication scores were computed using a series of Excel function commands (Appendix 10) to select specific cells if certain criteria were met, and aggregate the values, to obtain scores of node centrality, and edge weights. Figures 3.18 and 3.19, show the screenshots of one dataset - E1PreQID, and the rScript loaded into the RStudio environment for processing, and generating the csv output file. Figures 3.20, 3.21, and 3.22, show the output Excel file of E1PreQID. Content of columns is explained below.

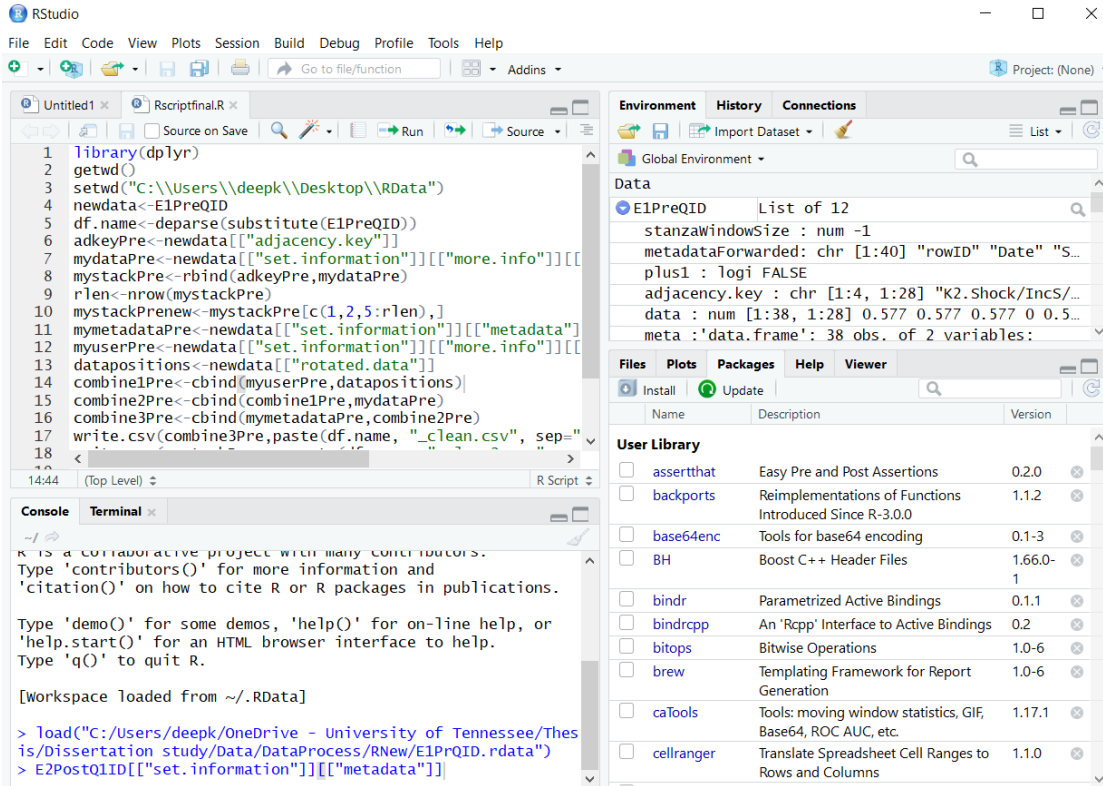


Figure 3.18. Screenshot shows the *rDataset* for PreQ of Experiment 1 (E1PreQ), loaded in the 'global environment' pane on the right, and the *rScript* to be run, loaded in the workspace pane on left.

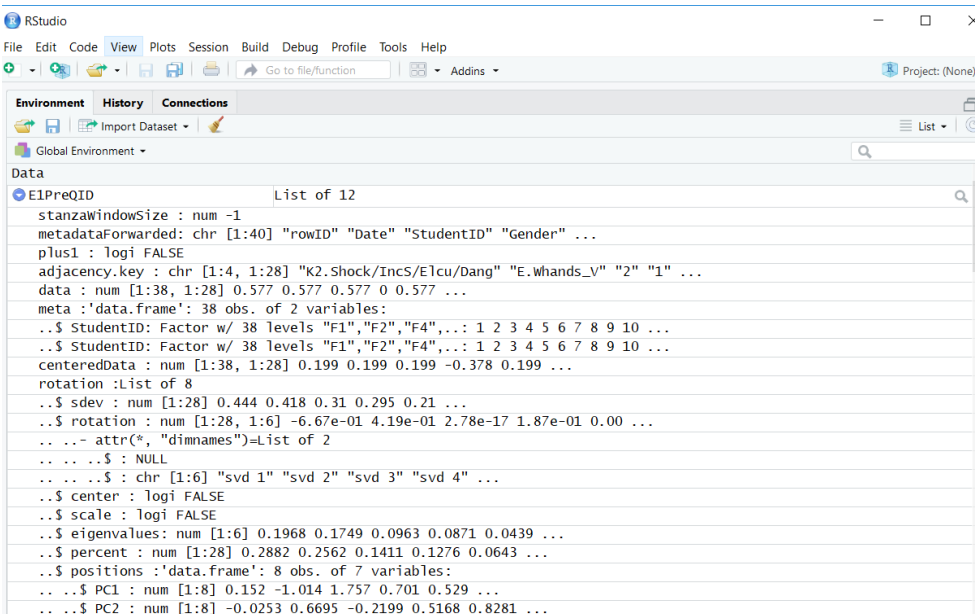


Figure 3.19. Screenshot shows the *rDataset* for PreQ of Experiment 1 (E1PreQ), loaded in the 'global environment' pane (Zoom view of the pane), with the list of 12 elements, which make up the dataset.

Date	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE
A	StudentID	Gender	Section	Studio	Expt	Q	axis_1	axis_2	axis_3	axis_4	axis_5	overall	EBAP5Gro	A3Attempt	A3(1)	GroQuiz	A3Group	Quiz4	Quiz4Gro	Lab4	Lab4Gro	H3Attempt	H3(1)	GroQuiz	H3Group	Test1	Test1Gro	Group	Username	Text
19-Sep F16	F	1	4	1	PreQ	1.7	1.5625	1.75	1	1	1.51667	Medium	6	Medium	8	High	3	Low	76	Low	8	High	8	High	9	Low	X1	@gptr1	#preyou	
19-Sep F41	F	3	4	1	PreQ	1.5	1.125	2	1.33333	2	1.5	Low	8	High	8	High	10	High	98	High	8	High	8	High	8	Low	J	@seiber	#preThe	
19-Sep M2	M	1	4	1	PreQ	1.1	2.375	1.5	1	2	1.66667	Medium	8	High	3	Low	81	Low	8	High	8	High	8	High	9	Low	F2	@cyohol	Distance	
19-Sep F25	F	3	4	1	PreQ	1.15	0.9375	2.125	0.33333	3.2	1.51667	Medium	8	High	8	High	8.5	Medium	92	Medium	8	High	8	High	13	High	O	@kostro	#preple	
19-Sep F31	F	2	4	1	PreQ	1.45	2	0.75	1.66667	1.6	1.75	High	6	Medium	7	Medium	10	High	96	Medium	8	High	8	High	13	High	L	@stapan	#preThe	
19-Sep F27	F	3	4	1	PreQ	1.25	1.875	1.75	1.33333	2	1.61667	Medium	7	Medium	8	High	7.5	Medium	101	High	8	High	8	High	6	Low	K2	@lentry	#preIf y	
19-Sep F10	F	1	4	1	PreQ	1.55	1.5	2.125	1.33333	1.4	1.63333	Medium	8	High	8	High	8	Medium	93	Medium	8	High	8	High	11	Medium	G	@coubbe	#pre1 you	
19-Sep F15	F	1	4	1	PreQ	1.6	1.5	2.5	1.33333	1.2	1.56667	Medium	8	High	8	High	7	Low	82	Low	7	Medium	8	High	12	Medium	A	@pre1 you	#pre1 you	
19-Sep F26	F	1	4	1	PreQ	1.45	1.375	1	1.66667	1.8	1.61667	Medium	7	Medium	8	High	8	Medium	98	High	8	High	8	High	5	Low	E	@kveime	If you tou	
19-Sep F38	F	1	4	1	PreQ	0.8	1.875	1.375	1	2.8	1.66667	Low	7	Medium	8	High	6	Low	82	Low	7	Medium	7	Medium	7	Low	A	@samed	#pre1 you	
19-Sep F14	F	1	4	1	PreQ	1.5	1.125	1.625	1.66667	2.6	1.65	Medium	8	High	8	High	8	Medium	88	Medium	8	High	8	High	10	Medium	H	@hamri	The volta	
19-Sep F36	F	1	4	1	PreQ	1.45	1.375	1.875	1.66667	2.8	1.7	High	7	Medium	8	High	10	High	96	Medium	8	High	8	High	10	Medium	C	@rwato	#pre1 the	
19-Sep F11	F	1	4	1	PreQ	1.5	1.375	1	1.33333	2.4	1.53333	Medium	2	Low	5	Low	8	Medium	92	Medium	2	Low	7	Medium	12	Medium	D1	@dreshi	You will f	
19-Sep M3	M	3	4	1	PreQ	1.55	1.125	2.875	1	2.8	1.76667	High	7	Medium	8	High	10	High	99	High	8	High	8	High	13	High	O	@fo128	Requestor	
19-Sep F29	F	3	4	1	PreQ	1.7	1.375	2	1	1.8	1.63333	Medium	8	High	8	High	8.5	Medium	103	High	8	High	8	High	13	High	R2	@Madisc	#pre1 The	
19-Sep F12	F	1	4	1	PreQ	1.5	1.4375	1.875	2	2.4	1.66667	Medium	8	High	8	High	6.5	Low	98	High	8	High	8	High	11	Medium	E	@ellis2	You woul	
19-Sep F23	F	1	4	1	PreQ	1.75	2	1.75	2.33333	0.8	1.71667	High	7	Medium	7	Medium	1	Low	81	Low	7	Medium	7	Medium	2	Low	F2	@kaylafi	#pre1 if y	
19-Sep F1	F	1	4	1	PreQ	1.7	1.9375	1	2.66667	1.4	1.71667	High	8	High	8	High	8	Medium	84	Low	8	High	8	High	7	Low	S	@lastua	#pre1 you	
19-Sep F21	F	1	4	1	PreQ	1.75	1.625	1.5	1.66667	1.4	1.61667	Medium	6	Medium	8	High	10	High	96	Medium	8	High	8	High	8	Low	L	@jezier	#pre1 The	
19-Sep F5	F	1	4	1	PreQ	1.85	0.875	2	2.33333	2.4	1.75	High	6	Medium	6	Low	4	Low	82	Low	8	High	8	High	9	Low	A	@athome	#pre1 if y	
19-Sep F20	F	1	4	1	PreQ	1.9	1.75	2.25	1	1.6	1.8	High	8	High	8	High	9	High	83	Low	8	High	8	High	9	Low	P	@_cooke	#pre1 you	

Figure 3.20 . Screenshot of columns A to AD of Output.

AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	BA	BB	BC	BD	BE	BF	BG	BH	BI	BJ	BK	BL	BM				
E.Wandt	K2.Shock	K2.V/V/Qz	K3.V/V/Qz	K2.WNCo	K2.WNCo	K1.wetnet	K1.WHBe	mysterPr	svd 1	svd 2	svd 3	svd 4	svd 5	svd 6	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19				
																			K2.Shock K2.V/V/Qz K2.V/V/Qz K3.V/V/Qz K3.V/V/Qz K2.WNCo K2.WNCo K2.WNCo K2.WNCo K2.WNCo K2.WNCo K1.wetnet K1.wetnet K1.wetnet K1.wetnet K1.wetnet K1.wetnet K1.wetnet K1.wetnet K1.wetnet																		
1	1	0	0	0	0	0	0	1	F23	-0.4828	-0.341	0.06179	-0.1164	-0.1122	0.0806	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
1	1	0	0	0	0	0	0	1	F1	-0.4828	-0.341	0.06179	-0.1164	-0.1122	0.0806	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	1	0	0	0	0	0	0	1	F21	-0.4828	-0.341	0.06179	-0.1164	-0.1122	0.0806	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	1	0	0	0	0	0	0	1	F30	-0.4828	-0.341	0.06179	-0.1164	-0.1122	0.0806	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	1	0	0	0	0	0	0	1	F9	-0.4828	-0.341	0.06179	-0.1164	-0.1122	0.0806	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	1	0	0	0	0	0	0	1	F2	-0.4828	-0.341	0.06179	-0.1164	-0.1122	0.0806	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	1	0	0	0	0	0	0	1	F25	-0.4828	-0.341	0.06179	-0.1164	-0.1122	0.0806	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
1	1	0	0	0	0	0	0	1	F7	-0.4828	-0.341	0.06179	-0.1164	-0.1122	0.0806	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
1	1	0	0	0	0	0	0	1	F19	-0.4828	-0.341	0.06179	-0.1164	-0.1122	0.0806	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
1	1	0	0	0	0	0	0	1	F44	-0.4828	-0.341	0.06179	-0.1164	-0.1122	0.0806	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
1	1	0	0	0	0	0	0	0	F57	-0.4252	0.31236	0.03962	0.10674	0.24485	-0.4203	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	1	0	0	0	0	0	0	0	F10	-0.4252	0.31236	0.03962	0.10674	0.24485	-0.4203	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	1	0	0	0	0	0	0	0	F25	-0.4252	0.31236	0.03962	0.10674	0.24485	-0.4203	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	1	0	0	0	0	0	0	1	F68	-0.2001	0.44788	-0.5359	0.37905	-0.0561	0.00721	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
1	1	0	0	0	0	0	0	1	F4	-0.2001	0.44788	-0.5359	0.37905	-0.0561	0.00721	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	1	0	0	0	0	0	0	1	M4	-0.2001	0.44788	-0.5359	0.37905	-0.0561	0.00721	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	1	0	0	0	0	0	0	1	F42	-0.2001	0.44788	-0.5359	0.37905	-0.0561	0.00721	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	1	0	0	0	0	0	0	0	F36	-0.0345	0.52849	0.50339	0.09025	0.1972	0.18459	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	1	0	0	0	0	0	0	0	F36	-0.0345	0.52849	0.50339	0.09025	0.1972	0.18459	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	1	0	0	0	0	0	0	0	F31	-0.0345	0.52849	0.50339	0.09025	0.1972	0.18459	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	1	0	0	0	0	0	0	0	M3	-0.0345	0.52849	0.50339	0.09025	0.1972	0.18459	0.57735	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	0	0	0	0	0	0	0	0	F12	0.06344	0.45447	0.34954	-0.0141	0.21672	0.25544	0.40825	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	0	0	0	0	0	0	0	1	F22	0.20701	0.39706	-0.6535	0.25206	-0.2686	0.42179	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	0	0	0	0	0	0	0	0	F36	0.24205	0.12511	-0.1122	-0.0977	-0.0236	0.01629	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
1	0	0	0	0	0	0	0	0	F25	0.38766	0.24742	-0.2945	-0.4648	0.12074	0.00289	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			

Figure 3.21. Screenshot of columns AF to AM of Output.

V20	V21	V22	V23	V24	V25	V26	V27	V28	CCS	TCS	I1CS	I2CS	S-level	K1C	K2C	K3C	
K1.wetneK1.wetneK1.WHBelK1.WHBelK1.WHBelK1.WHBelK1.WHBelK1.WHBelK1.WHBetterconductor																	
K2.WConK2.WNCoE.WhandK2.ShockK2.V//I/QvK3.V//I/QcK2.WConK2.WNCoK1.wetnesdecskin/BR									CCS	TCS	I1CS	I2CS		K1C	K2C	K3C	
0	0	0.57735	0.57735	0	0	0	0	0	0	0.57735	0.57735	0.57735	0	3	1.1547	1.1547	0
0	0	0.57735	0.57735	0	0	0	0	0	0	0.57735	0.57735	0.57735	0	3	1.1547	1.1547	0
0	0	0.57735	0.57735	0	0	0	0	0	0	0.57735	0.57735	0.57735	0	3	1.1547	1.1547	0
0	0	0.57735	0.57735	0	0	0	0	0	0	0.57735	0.57735	0.57735	0	3	1.1547	1.1547	0
0	0	0.57735	0.57735	0	0	0	0	0	0	0.57735	0.57735	0.57735	0	3	1.1547	1.1547	0
0	0	0.57735	0.57735	0	0	0	0	0	0	0.57735	0.57735	0.57735	0	3	1.1547	1.1547	0
0	0	0.57735	0.57735	0	0	0	0	0	0	0.57735	0.57735	0.57735	0	3	1.1547	1.1547	0
0	0	0.57735	0.57735	0	0	0	0	0	0	0.57735	0.57735	0.57735	0	3	1.1547	1.1547	0
0	0	0.57735	0.57735	0	0	0	0	0	0	0.57735	0.57735	0.57735	0	3	1.1547	1.1547	0
0	0	0.57735	0.57735	0	0	0	0	0	0	0.57735	0.57735	0.57735	0	3	1.1547	1.1547	0
0	0	0.57735	0.57735	0	0	0	0	0	0	0.57735	0.57735	0.57735	0	3	1.1547	1.1547	0
0	0	0.57735	0.57735	0	0	0	0	0	0	0.57735	0.57735	0.57735	0	3	1.1547	1.1547	0
0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	0	1	0
0	0	0	0	0	0	0	0	0	0	0.57735	0.57735	0.57735	0	3	1.1547	1.1547	0
0	0	0	0	0	0	0	0	0	0	0.57735	0.57735	0.57735	0	3	1.1547	1.1547	0
0	0	0	0	0	0	0	0	0	0	0.57735	0.57735	0.57735	0	3	1.1547	1.1547	0
0	0	0	0	0	0	0	0	0	0	0	0	2.3094	0	2	0	2.3094	0
0	0	0	0	0	0	0	0	0	0	0	0	2.3094	0	2	0	2.3094	0
0	0	0	0	0	0	0	0	0	0	0	0	2.3094	0	2	0	2.3094	0
0	0	0	0	0	0	0	0	0	0	0	0	2.3094	0	2	0	2.3094	0

In Figure 3.20 above, columns A to AD contain metadata of student scores on existing course assessments; column AE contains raw textual data on student response to that question. In Figure 3.21, columns AF to AM contain codes of occurrences of the code categories represented by these code columns, columns AO to AT contain the components or coordinates of positions of students' centroids along six dimensions of the projection space (svd1 to svd 6), and columns AU to BV (V1 to V28) contain values of strength of connections represented by each column, for each student, for that question. In Figure 3.22, columns CCS to I2CS contain the computed values of canonical connection strength or weight (CCS), transitional connection strength (TCS), intuitive 1, or concrete-intuitive connection strength (I1CS), and intuitive 2, or abstract-intuitive connection strength. S-level column contains the sophistication level for each student (each row) for that experiment, and the columns K1C to K3C contain computed values of centrality of K1, or canonical knowledge elements (K1C), centrality of K2, or concrete-intuitive knowledge elements (K2C), and centrality of K3, or abstract-intuitive knowledge elements (K3C). These measures of sophistication were computed as follows:

CCS - Sum of strength of connections between canonical knowledge elements (K1-K1) and between the concept constructed epistemically (E), and K1(E-K1).

TCS - Sum of strength of connections between knowledge elements when at least one of them was a canonical knowledge element (K1-K2, and K1-K3).

I1CS - Sum of strength of connections between concrete-intuitive knowledge elements (K2-K2) and between the concept constructed epistemically (E), and K2 (E-K2).

I2CS - Sum of strength of connections between abstract-intuitive knowledge elements (K3-K3), between the concept constructed epistemically (E), and K3 (E-K3), and between concrete and abstract-intuitive knowledge elements (K2-K3).

K1C - Sum of strength of connections which contained a canonical knowledge element, K1.

K2C - Sum of strength of connections which contained a concrete-intuitive knowledge element, K2.

K3C - Sum of strength of connections which contained an abstract-intuitive knowledge element, K3.

S-level - Sophistication levels were assigned based on the criteria explicated in Appendix 11.

The first four measures essentially gave an indication of strength of each type of connection in learners' thinking, and the next 3 measures indicated the strength of connections to each type of node or knowledge element in learners' thinking. The higher the values of CCs and K1C, the higher the level of

sophistication, and lower the values of K2C, K3C, I1CS, and I2CS, the higher the sophistication. TCS being a transitional measure, its strength indicated the transitory nature of learners' thinking, from intuitive or unsophisticated, to canonical, or sophisticated knowledge elements. These measures were analyzed statistically using SPSS, to answer RQ3. Descriptive statistics for each measure, bar charts of s-level, and correlational analyses with scores of other assessment instruments used in the course, were conducted.

Mixed Methods Analyses: Data Integration

The results of coding, ENA based assessment and RStudio, were consolidated through comparison, to identify the agreement and the disagreement between results. Findings were tabulated and graphed, for analysis, and to answer RQ4(a). the findings of correlational analyses were summarized, and Internal consistency reliability, Cronbach's alpha, was calculated, for the following sets of measures, to answer RQ4(b): (CCS, K1C, TCS), and (I1CS, I2CS, K2C, K3C). High degree of correlation was expected.

Validity and Reliability

Mixed methods. Since mixed methods studies involve both quantitative and qualitative methods, the validation of findings and research, is contentious at best. Johnson (2007) has developed a typology of mixed methods legitimization, by extending the integrative model of quality, design quality, and interpretive rigor, proposed by Teddlie and Tashakkori (2003, 2006): (i) Sample integration – to prevent legitimization issues arising from different sampling techniques of Quan and Qual; (ii) Inside-outside perspective of Qual and Quan (iii) sequential legitimization - influence of sequence on inference can be prevented by a multiple wave design (iv) conversion legitimization – need to avoid problems associated with verbal counting, acontextual counting, over-counting, by constructing narrative descriptions from quantitative data (v) paradigmatic mixing legitimization – while making meta-inferences researcher should evaluate the influence of his/her world view (vi) Commensurability legitimization – extent to which researcher is able to negotiate the cognitive switch between Qual and Quan worldviews (vii) Multiple validities legitimization – deciding the extent to which multiple validations are required, Qual, Quan and MM (viii) Political legitimization – arising from the differentiated power distribution of Qual - Quan, since power is centralized in the researcher, while it is placed in the hands of participants in qual.

The data collection of this study involved the same sample for both types of data, thereby addressing point number 1 above. As researcher, I recognized that power lay in the hands of participants, and that the validity of my findings rested on the participants' active participation; the study was conducted in multiple waves of activities over three studio sessions, and all results descriptively were integrated descriptively. The study also tested validity based on its relevance to the context of study;

thereby points 1, 3, 4, and 7, and 8 were addressed through design. Since as a researcher, I had experience in, and believed in the equal importance of both methodologies in education research, points 2, 5, 6 and 8 were also addressed.

Quantitative Methods

Of the different measures of validity and reliability described by Trochim (2006, 2015), this study ensured (i) Construct validity - the approximate truth of the conclusion that the operationalization proposed (sophistication of thinking) accurately reflected its construct (ii) Reliability – consistency or repeatability of measures (iii) Conclusion validity – degree to which conclusions reached were reasonable. External validity was addressed by describing the populations with which the study could be replicated in future (Chapter 5), based on proximal similarity with this population. Threats to conclusion validity were addressed by increasing reliability, maintaining a good implementation protocol, and making measurements in waves.

Construct validity was ensured and measured as follows: (i) Content validity- by checking the operationalization against relevant content domain, by giving a good detailed description of the content domain (ii) Face validity – by expert consensus that constructs meant what they said (iii) Concurrent validity is the ability to distinguish between groups – low correlation between learners with high and low sophistication, and medium correlation between low, medium, and high score groups of existing assessment measures was established (vi) Discriminant validity is degree to which the operationalization is not similar to (diverges from) other operationalization that it theoretically should be not be similar to – was established since low correlation was reported with Lab Reports, homework quiz, Lab Quiz, and Test1.

Internal consistency reliability was calculated across activities in a studio session, by calculating Cronbach's alpha, between equivalent measures of sophistication computed in this study, using the ENA assessment, and was found to be high. The study's design was developed, with validity and reliability in mind – hence the multiple phases, and correlational analyses with other tests such as EBAPS, quizzes, labs, and test.

Qualitative Methods

Per Trochim, et al., (2015), qualitative criteria of quality of research include: Credibility or internal validity; Transferability, or external validity; Dependability, or reliability; and Confirmability or objectivity. Credibility involves establishing that the results are believable from the participants' perspective (Bengtsson, 2016), by describing the data collection and analysis procedures in detail (Elo, and Kyngas, 2008). It can be enhanced by getting agreement from co-investigators, experts, or

participants themselves (Graneheim, and Lundman, 2004). In this study, two independent raters, one an undergraduate (in junior year) physics major, and the other, a Former Teaching Assistant of the course, reviewed the results of each qualitative analysis to ensure credibility. Inter-rater agreement was calculated, and disagreement among reviewers was resolved through discussion.

The sample of twitter messages selected for review was selected purposefully, rather than randomly, to represent a range of sophistication of thinking, by choosing a total of five participants - one from low score group, and two each from high and medium score groups of epistemic beliefs sophistication, as established through the EBAPS survey, for each question, and experiment. Each rater independently recorded her 'agreement' or 'disagreement' with the coding scheme developed by me, for each question, and experiment. So, for instance, if a response was assigned three or four codes by me, and the rater agreed to 2 out of 4, agreement was recorded as 0.5; if the rater agreed to 2 out of 3, then the agreement was recorded as 0.67, and so on. The inter-rater agreement table can be found in Appendix 15. The codes with disagreements were reviewed by me with the raters, and the final code was assigned through consensus.

Transferability refers to the degree to which qualitative results can be transferred to other contexts, or settings (Trochim, et. al., 2015). It is determined by how representative the sample is (Krippendorff, 2004). It can be enhanced by describing the context, criteria of selection, characteristics of participants, and data collection and analysis procedures (Graneheim, and Lundman, 2004). In this study, transferability was enhanced by maintaining a detailed description and log of all implementation protocols and procedures, with attention to any changes made. The sample used was also large enough (45 participants) to be representative of the population of undergraduate students taking PHYS 222 at University of Tennessee.

Dependability or reliability concerns with repeatability – whether the same results would be observed if the same 'thing' is observed twice (Trochim, et. al., 2015). It can be enhanced by keeping track of coding decisions, since they change over time (Bengtsson, 2016), and by linking results to data (Polit, and Beck, 2004). Implementation in multiple contexts of inquiry, enhanced the dependability of this study. Confirmability refers to the objectivity and neutrality of data, and the degree to which results could be corroborated by others (Bengtsson, 2016; Trochim, et.al., 2015). It can be enhanced through data audit, which examines research procedures, and makes judgements about potential bias (Trochim, et.al., 2015). In this study, two raters reviewed the results of qualitative analyses, and expert review process was used to refine the data collection instrument – question prompts. Thus, as discussed here, maximum efforts were made to achieve credibility, transferability, reliability, and objectivity.

Positionality. As a physics teacher with a passion for learning and teaching physics, I have experimented with numerous techniques and pedagogies during my teaching career of 17 years. My

teaching philosophy has evolved with experience, but I have retained my belief that a meaningful learning experience is central to the development of learners' positive attitudes towards physics. I have always held the bigger picture in mind wherein I see my students retaining the thinking skills that they develop in my classes, for an entire lifetime. While my philosophy framed the design of this study, I consciously bore in mind, the fact that I had taught physics in a different cultural environment, and that my own expectations and beliefs could influence the way I interpreted the response of students learning physics in a strikingly different cultural environment. Although I feel that I was immersed in the setting of the study, since I had also previously conducted a semester long qualitative inquiry about learners' perceptions of inquiry activities, in the same setting, through field observations, and open-ended survey, it is possible that when I analyzed student's thinking in this study, my expectations of teaching and learning derived from years of teaching physics, may have influenced my own thinking, to a certain extent.

Conclusion

This chapter began with a brief description of the research problem investigated through this study, and restate the research questions. A rationale for the single-study multi-phase mixed methods research design adopted in this study, was presented, along with a diagram which described the various phases of the implementation. Next, a description of context, participants, and research procedures was presented, followed by detailed description of quantitative data collection in phases 1 and 3, and qualitative data collection in phase 2, which also included a rationale for choice of Twitter as the primary data collection mechanism. I then elaborated the qualitative analyses, network analysis, quantitative analysis, and mixed methods procedures implemented. The chapter concluded with a discussion of validity and reliability protocols and measures examined for qualitative, quantitative, and mixed methods components of this study. In the next chapter, I will present the findings and results as they relate to the research questions of this study.

CHAPTER 4 FINDINGS AND RESULTS

Inquiry learning environments, which are commonly used in studio lab settings in introductory university physics courses, are designed to enable appropriate cognitive connections in learners' thinking, between the physical concepts being explored through inquiry, their mathematical, graphical and schematic representations, and the real world (Meltzer, and Thornton, 2012). However, existing assessment mechanisms are based on models of cognition and interpretation, which do not explicitly capture, characterize and monitor the development of, or enabling of 'connections' (NRC, 2001; Larkin, 2011). As discussed in Chapter 2, Epistemic Network Analysis (ENA) has emerged as an effective interpretive tool, in the context of virtual learning environments, to assess such connections, because it models learning as a network of connections between elements of knowledge, skills and/epistemologies, observed in learners' thinking (Arastoopour, et.al., 2012). The purpose of this study was to leverage four technologies – Twitter, Spreadsheet, ENA Web Tool, and RStudio, to develop and test the validity and reliability of an assessment framework, which utilized the lens of epistemic network analysis (ENA), to assess the sophistication of learners' thinking, during four inquiry experiments, conducted collaboratively, in studio lab setting of a non-Majors university physics course.

The social media platform – Twitter, was used to send question prompts to students, and to receive their responses before (PreQ), and after (PostQ) each experiment. Learners' Twitter responses were then exported to Spreadsheet, and coded qualitatively, to identify and categorize the knowledge elements in learners' framing of the following concepts: (a) relation between wetness/dryness of hands and their electrical resistance (Experiment 1) (b) change in capacitance of a parallel plate condenser (Experiment 2) (c) distinguishing ohmic and non-ohmic materials (Experiment 3) and (d) behavior of current in RC circuits (Experiment 4). After formatting the coded Spreadsheet data, to the specifications of ENA Web Tool 4.0, it was uploaded as separate CSV file, per question, to the ENA Web Tool. The ENA Tool modeled co-occurrences of code categories as connections between them, and transformed the co-occurrences into adjacency matrices, to create ENA datasets (rData), and plot the ENA sets as Network Graphs. These were then analyzed qualitatively and quantitatively (t-tests). The ENA sets (RData), were also exported to the statistical computing package RStudio, for further quantitative analyses – calculating and testing the reliability and validity of measures of node centrality and edge weight, of learners' epistemic networks. Specific research questions which guided the data collection and analyses were:

RQ1). What themes of 'epistemic framing' can be discerned from qualitative analyses of learners' Twitter responses, to question prompts sent to them before and after each activity, about the physical

concepts being investigated through each of the four inquiry experiments? How do the pre-and post-activity epistemic frames compare qualitatively? (QUAL)

RQ2). (a) What mean patterns of connections do the Network Graphs and Matrix Data reveal in learners' thinking before and after each of the four inquiry experiments (activities)? (b) Do the pre-and post-activity patterns differ significantly by (i) Section of course and (ii) Levels of (a) Epistemic beliefs Score (b) Lab Quiz Score (c) Lab Score, (d) Homework Quiz Score (Attempt1), and (e)Test1 Scores? (QUAL+QUAN)

RQ3). What levels of sophistication were identified in learners' thinking, through exploratory statistical analyses, of scores of node-centrality and weight density of connections? How do the measures of centrality and connection density of learners' Epistemic Networks relate to learners' (i) Lab Score (ii) Lab Quiz score, (iii) Test1score, (iv) EBAPS score, and (v) A3 HW Quiz score? (QUAN)

RQ4). (a) How do the inquiry activities compare with respect to their effect on the sophistication of learners' thinking? (b) Does the proposed ENA-based assessment framework provide a valid and reliable tool, to assess the sophistication of students' thinking in inquiry contexts in undergraduate physics courses? (MM)

This chapter will present findings related to the purpose of this study in five sections. First section will discuss the quantitative findings and results of EBAPS survey, and the grouping results of EBAPS, and four assessment instruments used in PHYS222 – A3 Home Work Quiz (Attempt1), Lab 4Quiz, Lab Report, and Test1. Second section will present the results of qualitative analyses of Twitter data – the broad themes, and code categories by question and Experiment, to answer RQ1. Third section will report findings of network analyses by Experiment and Question, to answer RQ2. Fourth section will answer RQ3, by drawing upon the results of Descriptive statistics, and correlational analyses, conducted using the ENA-based assessment scores of epistemic network sophistication, proposed in this study, and section five will answer RQ 4, by integrating the results of qualitative, quantitative and network analyses, using mixed methods.

Quantitative Results: Quantitative Instruments

Participant scores on Epistemological Beliefs about Physical Sciences (EBAPS) Survey, A3 Pre-Lab Home Work Quiz, In-Class Lab 4Quiz, Lab 4, and Mid-Term Test 1, were analyzed to group participants into three levels of performance by their scores on each of the five assessment instruments. These groupings, were later used to conduct statistical significance testing (t-tests) of differences in Mean Network Models by groups (reported in section 3 below). Individual scores, on these five assessment instruments were used for correlational analyses with scores derived using ENA (reported in section 4

below). This section presents the results of grouping, conducted using SPSS, by participant scores on four existing assessments used in the course, and the EBAPS survey. Overall results of EBAPS are also presented below.

Results of EBAPS Survey

Participants' responses to 30 items on the EBAPS survey instrument (which has been validated extensively through numerous studies, as discussed in Chapter 3) were scored as per the rubric appended in Appendix 3B. A total of 45 participants initially signed up for the study, and responded to the survey, but only 38 respondents participated in Twitter Activity during Studio Lab 4. Hence the analyses include data of 38 respondents. As seen from Table 4.1 below, not more than 15% of students chose Expert responses, to 30 Questions, categorized under five dimensions of beliefs.

Table 4.1 Percent Expert Responses on EBAPS Dimensions

EBAPS Axis or Dimension	EBAPS Survey Questions	% Expert Response
Structure of Knowledge	2, 8, 10, 15, 17, 19, 20, 23, 24, 28	14.3
Nature of Knowing (learning)	1, 7, 11, 12, 13, 18, 26, 30	10.5
Real life Applicability	3, 14, 19, 27	12.2
Evolving Knowledge	6, 28, 29	5.4
Source of Ability to learn	5, 9, 16, 22, 25	11.4

Results tabulated in Table 4.1 indicate that participants did not possess a sophisticated (expert-like) beliefs structure, about the structure of physics knowledge (memorizing facts, formulas, or a well-connected coherent structure), nature of learning physics (absorbing, or constructing understanding), real-life applicability of physics (interest limited to course or application to real-life), evolving nature of physics knowledge (evolving process from opinions, to evidence-based interpretations), and source of physics ability (natural or strategy, study, practice). The binary choices listed within parentheses in the previous statement, typify unsophisticated/novice-like beliefs and sophisticated/expert-like beliefs respectively.

Variable Grouping: Five Quantitative Instruments

Having confirmed from the results of EBAPS, that the undergraduates indeed possessed a relatively unsophisticated set of beliefs about physics, as was the premise of this study (Chapter 1 and 2), participant scores (variables) on four existing course assessments - A3Home Work Quiz (Attempt1) (Pre-Lab), Quiz 4 (In- Lab), Lab Report (In-Lab), Test 1 (Mid-Term), and EBAPS, were analyzed with the help of SPSS to group each of the five variables into three categories. Grouping was done to enable comparative qualitative Network analyses between groups, to answer RQ 2(b). Table 4.2 presents a summary of the statistics used for grouping.

Table 4.2 Descriptive Statistics for Overall EBAPS, A3Quiz (Attempt 1), Lab 4 Quiz, Lab 4, and Test1 scores

	A3 Quiz				
	Attempt1	Quiz4	Lab4	Test1	EBAPS
N	38	38	38	38	38
Mean	7.11	7.645	92.34	10.45	1.6342
Median	7.50	8.000	94.50	11.00	1.6500
Mode	8	10.0	96 ^a	11	1.5166
Std. Deviation	1.351	2.2478	7.722	2.748	.1269
Skewness	-2.351	-1.226	-.388	-.868	-.885
Minimum	2	1.0	76	2	1.1333
Maximum	8	10.0	103	14	1.8500
Percentiles					
10	5.90	3.900	81.90	6.90	1.3633
25	7.00	7.000	84.00	9.00	1.5125
33	7.00	7.000	88.00	9.87	1.5360
50	7.50	8.000	94.50	11.00	1.6500
66	8.00	8.500	98.00	11.74	1.7000
75	8.00	9.250	98.25	13.00	1.7166
90	8.00	10.000	101.20	14.00	1.7700

Note. a. Multiple modes exist, smallest values are shown.

Rationale. As discussed in Chapter 3, to test for concurrent validity of network measures assessed by ENA, it was necessary to identify other assessments which also measured sophistication of thinking and examine if students grouped by scores on these instruments, demonstrated significant differences in network characteristics (Trochim, et.al., 2006). Therefore, the goal of grouping was to define ‘high’, ‘medium’, and ‘low’ score groups for each dataset of scores on the five assessments identified. In other words, cutoff values had to be defined, such that the cutoff scores differentiated learners’ performance on these assessments. These five assessments were chosen, because by examining the ‘items’ which make up these assessments (Appendix 4A, 4B), it was evident that they measured

sophistication of thinking in terms of either conceptual understanding (A3Quiz, Quiz4, Test 1), problem solving (A3Quiz, Quiz4, Test 1), reasoning (Lab), and/ epistemic beliefs (EBAPS).

Student scores on each of these assessments are assigned using a rubric, and these scores, except EBAPS scores, count towards final grades. Grade boundaries and letters grades are assigned per Course policy, which is typically based on University grading scheme, and one way of sorting students for this study, could be to follow the norms established in the course. A grading model based on norms is called a norm-referenced grading model, and one of two methods – grading on the curve, or standard deviation method can be picked, to obtain the grade boundaries or cut-offs (Nitko, and Brookhart, 2007). Grading on the curve involves, selection of percentages arbitrarily, based on experience, as to what is realistic in that context, and thus requires an institution, or a Course instructor to make value judgements based on accumulated experience, and/previous assessment data, and is adopted by the University. Standard deviation method involves making decisions of grade boundaries, or cutoff scores, based on multiples of standard deviation, from the median score (50th percentile) – e.g. one standard deviation, or two standard deviations, and so on (Nitko, and Brookhart, 2007).

In this study, ‘curving on the grade’ was adopted to sort students into ‘high scores’, ‘medium scores’, and ‘low scores’, with ‘value judgement’ by me, to obtain maximum variation between groups, for each dataset of scores. The following process was adopted. (i) Score values corresponding to 33.33th percentile and 66.66th percentile were found. (ii) Values were rounded off to the first decimal for EBAPS, and nearest whole number for the other four datasets. (iii) If no scores fell below the 33rd percentile, purposeful decision was made as explained for A3Home Work Quiz (Attempt 1).

Statistical analysis. The following **commands were executed in SPSS.** (i) Analyze > Descriptive Statistics > Explore > Statistics > Percentiles, to obtain frequency distribution of scores (Table 4.2 above); (ii) Analyze > Descriptive Statistics > Frequencies > Statistics > Cut points (3 equal groups), to split each dataset of scores into three equal groups by percentile (33.33%) (Table 4.2).

EBAPS. For the 33.33th, and 66.66th percentile scores were 1.5360 and 1.7000 respectively, which were then rounded off to 1.5 and 1.7. Thus, scores less than or equal to 1.5 were categorized as ‘Low EBAPS’, scores greater than or equal to 1.7 were categorized as High EBAPS, and scores between 1.5 and 1.7 (excluding 1.5 and 1.7) were categorized as ‘Medium EBAPS’ (Table 4.3).

Table 4.3 Variable Grouping Table

Nominal Variable	Low	Medium	High	Max Score
EBAPS Group	≤ 1.5	$1.5 < \dots > 1.7$	≤ 1.7	4
A3Quiz (Attempt 1) Group	≤ 5	$= 6 \text{ or } 7$	8	8
Quiz 4 Group	≤ 7	$7 < \dots > 9$	$= 9 \text{ or } 10$	10
Lab4Group	≤ 87	$88 \leq \dots \leq 97$	≤ 98	100
Test 1 Group	≤ 9	$10 \leq \dots \leq 12$	≤ 14	15

Quiz 4, Lab4, and Test1. Same process was replicated for Quiz 4, Lab 4, and Test 1, and results are presented in Tables 4.2 and 4.3.

A3Quiz (Attempt1). For this variable, only two categories – Medium and High, could be statistically identified, since 50% of scores were in ‘High’ category. So, a purposeful decision was made, after examining Descriptive Statistics (Table 4.2), to use the 10th percentile as a reference point for deriving the ‘Low’ category. Table 4.3 above presents the results of grouping for all 5 variables.

Instructor Effect. Since the two sections were taught by two different Teaching Assistants, there could be an Instructor (Teaching Assistant) effect on observed Pre-Post activity differences in Epistemic Network Models by section of course. Hence, section of course was chosen as a categorical variable, to examine if indeed there was an instructor effect found. Second, by the theoretical framework which foregrounds this study, differences in thinking and framing of questions, and/differences in epistemic networks, could be explained by differences in prior beliefs and notions about the concepts, and about physics in general, because all students enrolled in the course were required to have fulfilled the same level of Math and Physics courses, and therefore, could be expected to possess the same repertoire of content knowledge and skills. To eliminate the possibility of a confounding effect of EBAPS, on epistemic network differences by Section, it was necessary to test the two samples (Section 1 and 3) for statistical significance of differences in EBAPS scores.

Table 4.4 Shapiro-Wilk Test of Normality

	Section	Statistic	Degree of freedom	Significance
Overall EBAPS	1	.904	29	.012*
	3	.976	17	.915*

Note. Null hypothesis for this Test is that the distribution is normal. *p < 0.05.

Results of test of normality (non-normal distribution), and non-parametric test of significance are summarized in Table 4.4 and Table 4.5 respectively. At a significance level of 0.05, no significant differences were found in EBAPS scores by section of course, since the significance value returned by the Mann-Whitney U Test was 0.316, which is greater than 0.05. (Table 4.5 below).

Table 4.5 Non-parametric Test for Independent Samples

Null Hypothesis	Test	Significance	Decision
The distribution of 'overall' is same across categories of 'Section'	Independent-Samples Mann-Whitney U Test	.316*	Retain null hypothesis

Note. Non-parametric test was chosen because the distribution of overall EBAPS scores was not normal for Section 1, while it was normal for Section 3, as revealed by the Shapiro-Wilk test at an alpha-level of 0.05 (Table 4.4). Shapiro-Wilk Test was chosen because N is less than 200. * $p < 0.05$.

Qualitative Results: Twitter responses

This section will present the major themes and categories identified through qualitative coding of learners' responses to Twitter prompts, sent to them before and after each of the four studio lab experiments, to answer Research Question 1: *"What themes of 'epistemic framing' can be discerned from qualitative analyses of learners' Twitter responses, to question prompts sent to them before and after each activity, about the physical concepts being investigated through each of the four inquiry experiments? How do the pre-and post-activity epistemic frames compare qualitatively?"* The Pre-Post Activity question prompts for each Experiment (Table 4.6) were sent as a Group Direct Message to the twitter IDs of all participants from my twitter handle, using TweetGuru MultiDM Tool for mass messaging. Pre-questions were sent before each Experiment, and Post-questions were sent 30 minutes later. This cycle was repeated for each Experiment, and each Experiment took about 40 minutes to complete. Each time, participants sent their responses as a Direct Message to my twitter ID. After each round of Pre-Post questions(Experiment), I exported the direct messages received, and sent from my account, to MS Spreadsheet, using the tool, Docteur-Tweety TwExList.

Table 4.6 Question Prompts

Experiment	Question Prompt
E1PreQ	What do you believe will happen if you touch the leads of a voltage source with wet hand? Why?
E1PostQ	What did you learn about the relationship between resistance of your body and its wetness, from this experiment?
E2PreQ	What are the ways by which you could change the capacitance of a parallel plate condenser? Elaborate.
E2PostQ1	What are the ways by which you changed capacitance in this experiment?
E2PostQ2	What did you learn about the relationship between capacitance of a parallel plate condenser and (a) distance between its plates and (b) area of the plates.
E3PreQ1	What are ohmic materials. Explain
E3PreQ2	What are non-ohmic materials. Explain.
E3PostQ1	Use your observations to decide if the bulb is an ohmic or a non-ohmic resistor. Explain your reasoning.
E3PostQ2	Use your observations to decide if the LED is an ohmic or a non-ohmic resistor. Explain your reasoning.
E4PreQ	Current in a circuit containing resistance R connected to a voltage source, reaches the value V/R almost instantaneously. How will the behavior of current change when a capacitor is introduced in the circuit to make it an RC circuit?
E4PostQ1	Describe what you found about the behavior of current in an RC circuit from this experiment?
E4PostQ2	Explain the similarity and difference you found in the behavior of voltage and current in an RC circuit.

Twitter Data for each Experiment was thus saved, coded, compiled, and formatted on a separate Excel file. Responses for each question were first processed and coded on a separate worksheet, and then merged together, along with other metadata (e.g. scores on various assessments and corresponding group assignment, Section, Gender) onto one worksheet. Deductive content analysis was used to code the content of each individual response (Twitter Direct Message), which formed the unit of qualitative analysis. Each row of Spreadsheet, thus represented one learner's (case) response to either a Pre-or post-question. As discussed in Chapter 3 (Methodology), I read each response, mentally identified the knowledge elements chosen by the student, and coded it almost verbatim, retaining the main content. Code columns were added as needed, and initial coding resulted in about twenty to thirty units of code per question. For each row i , and code column j , the cell (i, j) was assigned a value 1, if the response contained code j . Once all 38 responses for each question were coded, similar code categories were merged together through an iterative process, illustrated in Chapter 3. Frequency counts of each code category were used for further consolidation, and to eliminate code units with a count of less than 2. After all data was coded, I examined the code categories by question, and by Experiment, to identify common themes.

Themes of Epistemic Framing

In this sub-section I will elaborate the process of thematization. Specific codes and categories with examples, for each experiment and question, will be discussed in subsequent subsections, separately. All questions were designed to ask students to predict, describe, and/explain the behavior of physical entities or a physical event, in specified conditions. As I coded and consolidated codes for each question, within each experiment, I first had only two broad categories. One, the specific ‘behavior’ students were asked to describe or predict, e.g. behavior of ohmic and non-ohmic materials in Experiment 3, change in capacitance for experiment 2, behavior of current in RC circuits for Experiment 4, and holding a voltage source with wet hands, for experiment 1. Two, the knowledge elements used to construct or articulate the explanation, description, justification for the behavior specified in the question. I labelled codes in the first category as E. change in capacitance, for instance, to indicate that they were the physical behaviors or physical events framed, or constructed epistemically by learners (epistemically, because learners believed that they were formulating an appropriate construction of the said behavior). I labelled codes in the second category as K. follow ohm’s law, for instance, to indicate that that they were the knowledge elements used by learners to frame or epistemically construct the behavior or event. After plotting network graphs, and comparing graphs, I could ‘see’ that the knowledge elements were either ‘correct’ or ‘appropriate’, and ‘incorrect’, or ‘partially correct’, across all questions and experiments. To differentiate sophistication of connections on the network graphs, I had to further refine the category of ‘knowledge elements’.

So, I re-examined the codes in category ‘knowledge elements’ to differentiate it into two categories based on my judgement as teacher of physics, and as a researcher, who had reviewed the readings, and instructional materials assigned to students for each experiment - (i) correct or appropriate explanations, descriptions, justifications, and (ii) inappropriate, or partially appropriate responses, which seemed to have hints of intuition, or novice-like formulations, or prior misconceptions. I labelled codes in category (i) as ‘canonical knowledge elements (K1)’, and category (ii) as ‘intuitive knowledge elements (K2)’, based on the type of knowledge in terms of its canonical content. While K1 (appropriate knowledge elements), K2 (inappropriate, novice-like, partially correct, etc.) and E (behavior framed, or constructed epistemically or justified in reader’s view as appropriate) were clearly distinguishable, in terms of appropriateness, and in terms of the context of the question, and activity, K2 had many variations of ‘correctness’ or ‘appropriateness’. Researchers typically use classification reasoning, and their tacit and intuitive senses, to determine which data look and feel alike, for grouping them together (Saldana, 2009; Lincoln, and Guba, 1985, p.347). I decided, therefore, to look for ‘similarities’ and ‘differences’ between codes in category K2, from a teachers’ and researchers’ perspective.

I found that codes in category ‘intuitive knowledge elements’ differentiated in terms of origin – (a) concrete experience of inquiry experiment; (b) use of concrete (real life) entities instead of appropriate canonical physical quantities which are abstractions of those concrete entities; (c) partially correct understandings of course readings; (d) incorrect canonical relationships or abstractions; (e) hazy abstractions referring to indirectly related physical quantities. I then chose to draw upon Vygotsky’s (1962) classification of various stages of concept formation, based on the source or origin of learners’ experience, as discussed in Chapter 2, to further differentiate category K2. Since (a), (b), (c) above, had origins in learners’ concrete experiences, I labelled them ‘concrete-intuitive (K2)’. Some researchers have previously classified such knowledge as ‘intermediate’ or ‘transitional’ knowledge (Ernest, 2003) and treated it as a stage between canonical and intuitive, but I chose Vygotsky’s (1962) phases of concept formation, because his conceptualization of intuitive and canonical (conceptual, schooled or scientific) knowledge as mediating one another, rather than as a progression from intuitive to canonical, aligns well, with the basis of ‘co-occurrence’ of canonical and intuitive knowledge elements, which frames epistemic network analysis.

When an intuitive knowledge element could not be traced back to a concrete source or did not refer to a concrete entity, and seemed to have been abstracted by learners from their prior notions, beliefs, and/understandings which could be canonically appropriate or inappropriate, I classified it as ‘abstract-intuitive’ because its source happened to be learners’ own logic, or sense making of the underlying ‘physical mechanism’. Thus, code categories (d) and (e) above which originated in learners’ vague, partially correct, and/incorrect abstractions of prior schooled knowledge, were labelled as ‘abstract-intuitive (K3)’. diSessa’s (1988) p-prims, or phenomenological primitive model of knowledge, and Vygotsky’s (1962) second phase of concept formation, ‘complexes’, which are said to develop from learner’s concrete social experiences, subsume both themes I have identified, into one common category – intuitive knowledge. However, I argue, that differentiation of the category – intuitive knowledge, into two distinct categories, concrete-intuitive, and abstract-intuitive, based on the source of knowledge, provides instructionally valuable, and actionable information to an educator, who then can identify learner differences with specificity.

Thus, I finally identified four themes, which aptly capture the broad trends in students’ thinking. As described above, I chose to call them: **(i) Epistemic Construct (ii) Canonical physics (iii) Concrete-Intuitive physics and (iv) Abstract-Intuitive physics**. This thematic coding was informed by the broad categorization of learners’ knowledge structures as (a) canonical, scientific, and/expert knowledge and (b) intuitive, novice, naïve, and/primitive knowledge, advocated by both, conceptual change theorists

(diSessa, and Sherin), and epistemic frame theorists (Hammer, Redish), in the realm of physics education research. The four themes or trends found in learners' thinking, are summarized below.

Epistemic construct (EC). I chose to label this theme as 'Epistemic construct' (prefix E), because it represents and describes, the 'physics concept being framed' and constructed by learners, using one or more of the knowledge themes below.

Canonical physics (CP). For the population of students being examined, some explanations, descriptions, and/justifications, were found to be appropriate in terms of grade-level canonical 'physics' knowledge, in the context of studio lab 4, although distinctly different categories of knowledge representations – mathematical, graphical, verbal, etc., were used by the students, to frame each concept. This theme of epistemic frame was therefore labeled as 'canonical physics' (prefix K1). Responses assigned to this theme, typically made direct references to canonical physical quantities, such as 'distance between plates, area of a parallel plate capacitor and its capacitance', or 'wetness of hand and decrease in its electrical resistance'.

Concrete-Intuitive physics (IP-C). The prefix K2 was used to denote code categories falling under this theme. Responses assigned to this theme, demonstrated either an indirect, or somewhat vague association to canonical physical quantities normally attributed to the concept being framed, or referred to concrete, material entities, which embodied the canonical physical quantities associated with the concept. For instance, in Experiment 2, responses assigned to this theme, attributed 'decrease in capacitance' to 'increase in number of pages of the book', without making any reference to the canonical physical quantity, distance, which increased, when number of pages were increased, resulting in decrease of capacitance. Also, in this case, students' reference to 'number of pages' could be sourced back to 'course readings' before the experiment, and to 'students' observation logs' after the experiment.

Abstract-intuitive physics (IP-A). The prefix K3 was used to denote code categories falling under this theme. As discussed above, responses assigned to this theme, reflected learner's own understandings or constructions of 'physical concepts', abstracted incorrectly, from their prior knowledge, notions, and/beliefs. The label 'Abstract-intuitive Physics', is used to signify that the source of knowledge elements was learners' intuitive model or theory or 'abstraction' of the physical mechanism underlying the concept. The guiding principle of this theme, was 'canonically incorrect reasoning' or evidence of vital missing links in learner's model. In cases, where I found almost equal chance, that the source of knowledge could be either concrete or abstract, assignment was made to the higher theme of expertise of knowledge - concrete- intuitive.

Themes and specific code categories are elaborated for each Experiment separately below. Since the focus of qualitative analyses of this study, was identification of knowledge, skills, strategies, beliefs pertaining to grade level physics, the code labels chosen, sought to retain ‘physics language’ verbatim, so that richness of information was not lost; and theme labels sought to characterize the level of physics expertise, without quantizing it. While these analyses were informed by other works in the field of physics education research, which subscribe to the resource-based, or knowledge-in-pieces model of cognition, and which characterize learners’ thinking on the basis of qualitative interpretation of learners’ responses alone (Tuminaro, 2004; Bing, 2008) the goal, was not to characterize thinking, but to provide ‘coded qualitative data’ which was then further analyzed using network analyses, to draw mixed methods inferences - qualitative as well as quantitative, about the sophistication of learners’ thinking.

Qualitative Findings: Experiment 1 (The Resistance of Your Body)

In this experiment, students measured the resistance of their body with dry and wet hands and predicted the current flowing through their body in each instance. Specific code categories identified under each theme, by question, are presented below (Table 4.7). Students in this Course are provided online access to class notes, OpenStax Text Book, and Studio Labs. They are expected to read the theory and concepts associated with the Labs and take a multiple-choice homework quiz which allows multiple attempts, ahead of class. The reading material for Experiment 1 (Chapter 20-21, Textbook), included a section on – Electric Hazards and the Human Body, which clearly outlines the effects of electric current on human body, for various ranges of current, and explains how wetness decreases the resistance of body, because of ions present in water.

Yet, the Twitter **pre-question** for this experiment, “what will happen if you touch the leads of a voltage source with wet hands”, triggered a response, “receive shock”, from 23 participants. It was classified as ‘**concrete-Intuitive**’, because (i) canonically speaking, such a conclusion would require other information (not provided in the question) such as - the voltage of source and resistance of wet hands, and (ii) students’ intuitive knowledge could be sourced to ‘prior readings’. 22 participants gave a canonical response, referring correctly, to “increase in conductance” (17 students), or “decrease in resistance” (5 responses) of hands, skin/body, when wet, and these code categories were therefore classified as ‘**canonical**’ knowledge elements.

Table 4.7 Experiment 1 Code Categories and Themes

Question	Theme	Code Category	Description	Frequency
PreQ	Epistemic Construct	E.Whands_V	Voltage source in wet hands	38
		Canonical	K1.wetnessdecskin/BR	Wetness decreases skin/body resistance
			K1.WHBetterconductor	Wet hands better conductor
		Concrete Intuitive	K2.WConductor	Water is a conductor
			K2.WNConductor	Water is not a conductor
			K2.Shock/IncS/Elcu/Dang	Shock/Increased shock/Electrocution
			K2.V/I/Qwlinec	Voltage/Current/Charge will increase
		Abstract-Intuitive	K3.V/I/Qdec	Voltage/Current/Charge will decrease
PostQ	Epistemic Construct	E.RelWetness,R	Relation between wetness of hand and resistance	38
		Canonical	K1.WetnessdecR	Wetness decrease resistance
			K1.WetnessincCond/I	wetness increases conductance/current
			K1.dryHHhigherR	Dry hands, Higher resistance
			K1.I,RvarywWetness	Current, resistance vary with wetness
		Concrete Intuitive	K2.Noshock	No shock
		Abstract-Intuitive	K3.WetnessINcR,decCond	Wetness increases R, decreases conductance

The response “increase in voltage, current, and/ charge” given by 9 participants, was also classified as ‘**concrete-intuitive**’, for two reasons, One, there was no direct reference to canonical relationship between resistance of hand, and wetness. Two, the indirect reference to increase in voltage, current/charge, which would accompany the decrease in resistance of hand due to wetness, could be interpreted in one of two ways - either as students’ frame being directed in the direction of canonical knowledge, or it could also be, that their abstraction of ‘increase in voltage, current and/charge’, was associated with the introduction of ‘voltage source’. The latter interpretation would classify the code category as ‘abstract intuitive’. But by the criteria chosen to be followed, in case of equal probability of both events (themes), assignment was made to the theme higher up in the expert-novice hierarchy – ‘concrete-intuitive’ in this case. Similarly, responses “Water is a conductor” (9 participants) and “water is not a conductor” (2 participants) were classified as ‘**concrete-intuitive**’ in the context of this question, because canonically, such a conclusion would require information about purity (pure water by itself will

not conduct current) of water, and/ concentration of salt ions present in water - information which was not provided in the question. Students were referring to canonical ideas from readings, intuitively.

4 participants' epistemic frame of "a voltage source held with wet hands", consisted of knowledge elements which were coded as "voltage/current/charge will decrease", and classified as '**abstract-intuitive** physics', because they reflected an abstraction, based on prior misbeliefs or notions underlying the concept framed, since what happens is, 'an increase', not 'a decrease', of these physical quantities. Overall, there were more responses that contained 'concrete-intuitive physics' (43), than responses which reflected 'canonical physics' (22), before the Experiment.

After Experiment 1, students were sent the question, "What did you learn about the relationship between resistance of your body and its wetness, from this experiment?" (PostQ). 49 students responded **canonically**, while only 4 participants' epistemic frames were characterized as **Intuitive** (Table 4.7). This suggested that the Experiment was productive in terms of redirecting learners' intuitive notions about the relationship between wetness of hand and its electrical resistance. However, it was also evident from the analyses, that for PreQ, the three themes did not occur alone in learner's response, and quite often, canonical and intuitive knowledge elements would occur together. For instance, student F2's response to PreQ was "... *they should feel the shock of that wire because wet hands are providing conducting path to the current to flow inside the live body from electrical wire ...*". In this response, canonical element "wet hands better conductor" co-occurs with intuitive element, "feel shock", and when analyzed (counted) as a connection or co-occurrence, as will be done through network analysis presented in the next section, rather than counting the occurrences separately, more information becomes available with respect to learners' framing of the concept.

Response of F33 to Post Q: "*The resistance in our body decreased when our fingers were wet as I expected. However, we did not get shocked like I originally thought we might.*", exemplifies partial resolution of prior intuitive idea, through inquiry, since the student seemed to have preserved unreconciled notion about 'getting shock' - unreconciled, because she still did not connect to the knowledge element that 'threshold values of voltage and resistance must be crossed, for shock to be experienced by the human body'. Although the focus of the experiment was 'wetness and resistance', significant number of students, like F33, seemed to frame the concept in terms of shock before the experiment, and this 'node (Knowledge Element)' did not get the opportunity to get associated with or connected to canonical knowledge element mentioned in the last statement, while conducting the Experiment.

Such opportunity could be provided by the TAs explicitly talking about it, when such ‘hanging nodes’ are found to dominate learners’ thinking. However, existing assessments are not equipped to identify and/ capture the ‘hanging nodes and/connections’ which get uncovered through network analyses. The other reason for intuitive framing in terms of ‘shock’, could be that the student population taking this course have biology related majors, and therefore value or give more weight to ‘effects on the human body’, and think deeply about them, as compared to students majoring in Engineering or physics. While the experiment does seek to build upon what they would ‘value’ - by providing opportunity to relate physical concepts with the human body, the opportunity remains underutilized, for want of a mechanism to track students’ thinking, identify ‘hanging nodes and connections’, and provide ‘anchors’ to sophisticate learners’ frames.

Qualitative Findings: Experiment 2 (Measuring Capacitance)

In this experiment, students constructed a variable ‘parallel plate capacitor’ of sorts, by using two rectangular metal foils (having handles to insert the probes of a multimeter), separated by the pages of a book, and investigated how the ‘capacitance’, C , of a parallel plate capacitor, changed, when the area, A , of the plates was changed (keeping distance constant), and when the distance, d , between the plates was changed (keeping area constant). They measured capacitance using multimeter, for different number of pages of the book, between the plates (changing distance), and for different areas of the plates (changing overlap of plates between the pages of book, to half, and then one fourth, of the original overlap). They also used Excel to plot and interpret the relationships between C , A , and d , graphically. Thus, the inquiry was expected to enable connections between mathematical, graphical, and verbal representations of physical relationships governing ‘capacitance’. Before the studio session, students were also supposed to review class notes for the studio session, and the corresponding content from Text Book Chapter 19, on ‘Capacitors and Dielectrics’, explaining the physical ideas underlying the concept of a capacitor and a parallel plate capacitor.

The Twitter PreQ for this Experiment, asked students to explain how they could change the capacitance of a parallel plate condenser, PostQ1asked them to explain how they changed the capacitance in the experiment, and PostQ2 asked them to draw conclusions about the relationship between capacitance, area and distance, based on their observations. Specific code categories identified from students’ responses, under each theme, by question, are presented in Table 4.8.

Before the experiment, out of 44 students who used canonical physics to describe the concept of change in capacitance, 29 students framed it in terms of change in distance, and 15 students, framed it

in terms of change in area. Latter 15 responses mostly co-occurred with former 29 responses, which suggested that 14 (= 29-15) students framed change in capacitance, only in terms of distance, and not area. 6 responses which refer to number of pages and overlap area, as ways of changing capacitance, were classified as **'concrete intuitive'**, because while students did not further connect the 'number of pages' and 'overlap area' to canonical physical quantities of 'distance' and 'area' respectively, their intuitive response could be sourced back to 'prior readings of the studio session', which informed students that they would be changing capacitance by changing the number of pages and overlap area. However, thinking of change in capacitance in terms of change of 'concrete' entities, with no evidence presented, of abstraction to the canonical physical quantities being changed as a result, reflected superficial, not conceptual thinking, and hence the classification as **'concrete intuitive'**.

18 responses made vague (not concrete) references to physical quantities not completely out of the realm of quantities which could be construed as being associated with change in capacitance of a parallel plate capacitor –(i) weight, size, shape, orientation, of circuit or capacitor, and (ii) voltage (V), current(I), charge(Q), or resistance(R). In a physical sense, one or more of quantities in (i) could effect change in area - size and shape directly, and weight or mass indirectly, but the response was classified as **'abstract intuitive'**, since there was no evidence that students made the connection to canonical quantities of area, and or distance, or that their intuitive knowledge originated from a concrete source in this context, such as 'reading materials'. Similarly, V, I, Q, and R (association with C, more far-fetched than V, I and Q), in (ii), could change because of change in capacitance, but cannot be said to cause a change in capacitance, suggesting that students' abstraction of the causal relationship was mostly intuitive.

Thus, responses in (i) and (ii) (above paragraph) were classified as **abstract-intuitive**, because they were interpreted as originating from students' seemingly logical abstraction, of prior understandings of these physical quantities, and of capacitance, and a belief that (i) capacitance could be changed by changing one or more of these quantities, either with respect to circuit or capacitor, or (ii) change in capacitance and change in V, I, R, and Q, are somehow related. The reason for reference to circuit, remained unclear, but it could be, that students were thinking of a 'parallel plate capacitor assembly' as a circuit in and of itself (laying the foundation, perhaps, of the development of a misconception of 'capacitor', if not identified and redirected right away).

Table 4.8 Experiment 2 Code Categories and Themes

Question	Theme	Code Category	Description	Count
PreQ	Epistemic Construct	E.Expect_ChangeC	Ways to change Capacitance	46
		K1.ChangeD	Change Distance between plates	29
		K1.ChangeA	Change Area of plates	15
		K2.ChangeNP	Change number of pages	4
		K2.ChangeOA	Change overlap area	2
		K3.ChangeCct/Plsi/sh/ori/W	Change circuit/plate size, shape, weight or orientation	8
PostQ1	Epistemic Construct	K3.ChangeI/V/R/Q	Change current/voltage/resistance/charge	10
		E.Activity_ChangeC	Ways Capacitance was changed	44
		K1.changeD/thickness	Changed distance between plates	20
		K1.changeA	Changed area of plates	8
		K2.changeNPgs	Changed number of pages	21
		K2.changeOverlap	Changed overlap area	4
PostQ2	Epistemic Construct	K2.sepFoilabyggs	Separated foils by pages	2
		K3.changeI/V/R/Q	Changed current/voltage/resistance/charge	2
		E.Interpret_RelC,A,D	Relationship between C, A, D	40
		K1.CdecwIncD	C decreases with increase in distance	27
		K1.CIncwA	C increases with area	10
		K1.CincwDecD	C increases with decrease in distance	4
		K1.C,dinvrel	C and d have inverse relationship	2
		K1.C,AdirRel	C and A have direct relationship	2
		K1.CdecwA	C decreases with area	8
		K2.CdecwIncNP	C decreases with increase in no of pages	6
		K3.CN/lchangewA	C does not change with area	2
		K3.R/VincwIncA	Resistance/Voltage increases with area	4
		K3.R/V/WdecwIncD	Resistance/voltage/weight decreases with increase in distance	6
		K3.CincwdecOA	C increases with decrease in overlap area	2
		K3.CdecwIncOA	C decreases with increase in overlap area	2

In response to Post Q1 which asked students how they had changed capacitance in the experiment, 28 students responded **canonically** after the experiment, as compared to 29 before experiment; 25 responded '**concrete intuitively**' based on the observations taken during the experiment, in comparison to 6, before the experiment; and only 2 responded '**abstract intuitively**', in term of changing V, I, R, and/Q, as opposed to 18 before the experiment. Thus, when I looked at PostQ1 responses alone, the experiment seemed to have been effective in shifting the focus of 16 students, at the very least, from abstract intuitive formulations, to concrete entities associated with change in capacitance, viz., number of pages, and overlap area. However, on looking at responses to **PostQ2**, which asked students to explicitly make conclusions based on their concrete experience, about the relationship between C, A, and D, I found that - (i) Less number of students (18) could describe verbally, the causal relationship between C and A, while 27 correctly described it for D, which could mean, that students could 'see' the relation between C and D, but not connect to the relation between C and A, through the inquiry experiment. These were classified as **Canonical** KEs. (ii) 6 students described the relationship in terms of a concrete entity, 'number of pages' of the book, instead of the canonical quantity, distance, hence the KE was classified as **Concrete Intuitive**. (iii) 6 students were not able to abstract the correct causal relation between Capacitance, and area or between overlap area and C (4 students), while 10 others, abstracted relationships in terms of Resistance, Voltage, or weight, instead of C. The corresponding KEs were classified as **abstract-intuitive**.

In conclusion, approximately 18 students' frame did change from abstract-intuitive before experiment, to concrete-intuitive after the experiment, but they still were unable to draw correct inferences or abstract correctly to canonical causal relationships, after the experiment. Besides, out of 45 students who inferred canonically for PostQ 2, only 10 referred to area. These students would clearly benefit from TA's redirection of their attention to canonically relevant quantities and relationships, but existing assessments do not provide opportunities to the TAs, to obtain such information about learners' missed connections. It would be useful to the TAs and the learners, if learner's thinking could be expressed, and monitored before, during and after activities, to enable administration of instructional remedy in real time.

Qualitative Findings: Experiment 3 (Ohmic and Non-ohmic Materials)

In this experiment, students investigated the relationships between voltage and current, in ohmic and non-ohmic conductors. They used a computer-aided Pasco RLC (resistance, inductance, capacitance) Board Interface, two ohmic conductors - 33 ohms and 100 ohms resistors, and two non-ohmic conductors - light bulb, and LED (light emitting diode), to record graphically, the relation between voltage (V) across each of these circuit elements, and current (I) flowing through them, for specified values and types of

input voltage (DC or pulse). PreQ1 asked them to explain “What are ohmic materials”, and PreQ2 asked them to explain “What are non-ohmic materials”. PostQ1 asked them to “Use your observations to decide if the bulb is an ohmic or a non-ohmic resistor, and explain your reasoning”, and PostQ2 asked them to “Use your observations to decide if the LED is an ohmic or a non-ohmic resistor, and explain your reasoning”. Post questions, were like questions of their Inquiry Labs. Course readings for this experiment, included sections of the Textbook (Chapter 21) covering ‘Resistors in series and parallel’. Specific code categories identified from students’ responses, under each theme, by question, are presented below (Table 4.9).

Before experiment, 32 out of 43 descriptions of ohmic materials were canonical – 8 descriptions used graphical representation, 16 students described the relationship between voltage and current, or resistance, and 8 responses did not elaborate on their conceptualization of Ohm’s law, although their response that ohmic materials ‘follow ohm’s law’, were appropriate. 6 responses characterized ohmic materials vaguely as ‘linear’ or as ‘act as resistor in circuit’. Since they did not specify clearly, the canonical relationships associate with ohmic materials, but did seem informed either by course readings, or by prior understandings, they were classified as ‘concrete-intuitive’. Three responses which referred to ‘current remaining constant’ clearly seemed to be abstracted from prior misbeliefs or notions about ohms’ law, according to which the ratio of voltage to current remains a constant, and not current. Hence, they were classified as ‘abstract-intuitive’. Non-ohmic materials were described canonically in 28 out of 37 responses – 8 of which used graphical representations, 12 described in terms of canonical relationships between physical quantities, and 8 merely stated correctly, that these materials did not follow ohm’s law. Three responses made vague reference to non-linearity of non-ohmic materials, without specifying canonical relationships, and therefore were classified as concrete-intuitive.

Out of 4 responses classified as abstract-intuitive, 2 responses referred to ‘current not being constant’, and 2 others, referred to ‘there being no voltage or current’. These abstractions seemed to originate in learners’ prior ‘undefined’ beliefs or misbeliefs about the causal or associational relationship between the ohmic or non-ohmic nature of a circuit element, and current flowing through it, or voltage across it. For instance, students could be associating their knowledge that the ratio of voltage to current, or resistance does not remain constant for non-ohmic materials, with the other quantities associated with resistance, viz. voltage, and/current. Hence the classification as abstract-intuitive.

Table 4.9 Experiment 3 Code Categories and Themes

Question	Theme	Code Category	Description	Count
PreQ1	Epistemic Construct	E.Ohmic	Ohmic Materials	43
		Canonical	K1.Const R k1.FLOhmL K1.LinGraph K1.LinVIRel K1.StLineGraph	9 8 4 7 4
	Concrete-Intuitive	K2.Linear	Are linear	4
		K2.ActasRinCct	Act as resistance in circuit	2
	Abstract-Intuitive	K3.Const I	Current remains constant	3
PreQ2	Epistemic Construct	E.NonOhmic	Non-ohmic Materials	37
		Canonical	K1.DNFLOhm K1.NLinVIRel K1.NConstantR K1.NlinVIGraph K1.Nstlinegraph	8 6 6 4 4
	Concrete-Intuitive	K2.Nlinear	Are non-linear	3
	Abstract-Intuitive	K3.NIorV	Have no current or voltage	2
		K3.NotConstantI	Current is not constant	2
PostQ1	Epistemic Construct	E.Ohmic	Ohmic	11
		E.Nohmic	Non-Ohmic	29
	Canonical	K1.NConstR	Resistance was not constant	5
		K1.NFOhm	Did not follow Ohm's law	2
		K1.NLinG/NconstS	Graph was not linear/slope was not constant	17
	Concrete-Intuitive	K2.NLRel/freq/NConstflow	Relation or frequency was nonlinear/ current flow was not constant	4
		K2.INreltoV	Current was not related to voltage	4
	Abstract-Intuitive	K3.ConstR K3.LinS/G	Resistance was constant Graph and/slope was linear	2 7
PostQ2	Epistemic Construct	E.Ohmic	Ohmic	8
		E.Nohmic	Non-Ohmic	34
	Canonical	K1.NConstR	Resistance was not constant	5
		K1.Nlin/NConS	Slope was nonlinear/not constant	16
		K1.NLinVirel	Voltage, current relationship was not linear	3
	Concrete-Intuitive	K2.NLinear	Were not linear	3
		K2.LessConsistLinGraph	Had less consistent linear graph	2
	Abstract-Intuitive	K3.LinG/Pts K3.linear K3.O/NORElto I	Had linear graph/points Were linear Ohmic or non-ohmic behavior was related to current	2 1 2

Although most students seemed to have at least basic canonical understanding about classification of materials as ohmic, or non-ohmic, before the experiment, their responses after the experiment indicated that in the context of applying the knowledge to make inferences, based on graphs, many were ‘confused’ by the non-linear graphs, possibly, because they encountered these graphs in practice, for the first time, and what they saw was not what they ‘expected’.

Post Q1. Out of 11 students, who identified light bulb incorrectly, as ‘ohmic’, 9 students justified their answer, in terms of canonical knowledge of ohmic materials – constant resistance (slope) and linear graph, when the graph was in fact not linear, as was confirmed by their lab report data. Their prior ideas about linear and non-linear graphs, or possibly, lack of concrete prior experience, seemed to be hindering correct inference of their observations.

For instance, F10 responded “*the bulb is slightly ohmic because the graph was not completely linear*” Some other responses were - F14: “*The bulb is ohmic because the graph we produced is linearly increasing*”; F21: “*we decided the regular bulb was almost an ohmic resistor because the graphic relationship was nearly linear*”. When I cross-checked with copies of their lab reports, I found that they did have difficulty in ‘reconciling’ the graph as ‘non-linear’. F10, F14 and F21 lost points on this specific question of Lab4, since these answers were replicated on the labs as well. These 9 responses were therefore classified as ‘**abstract-intuitive**’. This finding lends support to the argument, that while TA comments and feedback received one week later, maybe useful in knowing the right answer, the same feedback could yield more productive outcomes in terms of students’ epistemic construction, if given in-context.

24 out of 29 students who correctly identified bulb as non-ohmic, gave canonically correct justifications – 17 of them used graphical knowledge representations and 7 used canonical relationships between V, I, R. These code categories were classified as ‘**canonical**’ knowledge. Out of 8 ‘**concrete-intuitive**’ responses, 4 referred vaguely to voltage and current being ‘not related’, and 4 referred to non-linear relationship, non-linear frequency, and/ non-constant flow of current. These responses suggest that while these students decided that the behavior of light bulb was non-ohmic, they had difficulty in associating their concrete observations with canonical relationships between relevant physical quantities, and canonical terms. Post Q2 responses for LED, mostly mirrored the results of Post Q1. Differences are discussed below.

Post Q2. 34 students correctly identified LED as non-ohmic. 16 of them reasoned that the graphs were non-linear, and/ slopes were not constant. 8 students abstracted their reasoning to non-linear relationship between voltage, and current, and / non-constant resistance. These responses were classified as ‘**canonical**’ for referring to appropriate canonical knowledge. 3 responses which referred vaguely, to

the LED being non-linear, and 2 responses which referred to the graphs as being ‘less consistently linear’, were classified as ‘**concrete-intuitive**’, because they reflected students’ difficulty in attributing concrete observations, to corresponding canonical knowledge – specifying graph or voltage-current quantities in the first instance (nonlinear), and non-linearity of the graph, in the second instance (less consistent linear). 3 students who misread the graphs as linear, and 2, who concluded that ohmic and non-ohmic behavior depended on the value of current, were classified as ‘**abstract intuitive**’ since their reasoning seemed to be influenced by lack of prior experience with making inferences about real-time graphical outputs, which sometimes are also accompanied by ‘background noise’ depending on the hardware settings chosen by users. Example ‘abstract-intuitive’ responses – F29: *“I would say that LED is both an ohmic and non ohmic. At a certain current it become ohmic but the rest of the time it is non ohmic; F16: “According to our data the LED was Ohmic because it did display linear points”.*

To sum up, on an average, 25% of the students (10 to 11, out of 41 to 42 responses to Post Q1 and Post Q2 respectively) seemed to have ‘prior expectations’ about ‘neatly defined’ graphical features, which got in the way of their making sense of non-linear graphical VI outputs generated by light bulb, and LED, as non-linear, without attributing some ‘linearity’ to it, based on certain features, which conflicted with their expectations of non-linearity. However, more than half of these 25% percent students, (approximately 88% of total respondents to PreQ1 and PreQ2) did seem to have canonical understanding that ohmic materials would demonstrate linear VI relationships and non-ohmic materials would demonstrate non-linear VI relationships, before the experiment. Second, the non-linear graphical output of bulb was more confusing to students, than the LED. While the activity’s goal might have been to enable connections between theory and real life, it remained underachieved, because new ‘epistemological issues’ arose in the minds of some students, with respect to electrical properties of bulb and LED, which remained unresolved, for want of systematic mechanisms (other than TA’s attention when asked for), to capture ‘epistemological difficulties’ in real time, and address them in-context. Although the numbers tell a story of overall success of the experiment, these qualitative findings do tell a story of ‘hidden epistemological issues’ which need to be attended to, if we wish to address effectively, the broader problem of low participation and interest in physics.

Qualitative Findings: Experiment 4 (RC Circuits)

In this experiment, students used the Pasco 850 interface to investigate the behavior of current and voltage across a capacitor in an RC circuit, by using a square wave input signal, and recording the graphical output for three seconds. They were asked to use the graphs, to find and compare time constants of current, voltage, and the mathematically calculated value using the formula. Since RC circuits are the most common types of basic circuits which underlie the ‘timing mechanisms’ of a wide range of

electronic gadgets, such as the flash of a camera, or are instrumental in defining the bandwidth in telecommunication contexts, it is important that students develop a ‘scientific mental model’ of how they work. Not unlike the previous experiment, the critical component of ‘getting it right’, was ‘interpreting the graph right’. Students had online access to the lab, class notes and textbook content, pertaining to the working of RC circuits. Themes and code categories by question, are discussed below, with reference to Table 4.10.

PreQ asked students to explain how the behavior of current in a circuit containing only resistance, would change, when a capacitor was introduced in series with the resistor. Out of 39 responses, 26 students used canonical language, and explained, that current would decrease, circuit would switch on-off at intervals, time taken by current to reach the peak value would increase, capacitor would store energy, or that the current would stabilize when capacitor got charged. Since one or more of these explanations were canonically reasonable, they were classified as **canonical**. While 3 students made correct reference to ‘change in voltage and/ current’, and 2 students predicted that voltage would decrease, their responses were deemed vague and therefore, intuitive, because if the change in current were ‘an increase’ it would be incorrect, and student’s reasoning model of associating decrease in voltage with change in current, was not specified. These 5 responses were hence classified as **concrete-intuitive**. 7 responses, which predicted changes in a direction opposite to the canonical change, seemed to be influenced by prior misconception about the physical mechanism underlying the charging and discharging of a capacitor, which would, cause the changes in current and/voltage. Hence, they were classified as **abstract-intuitive**.

PostQ1 asked students to explain what they had learned about the behavior of current in RC circuits, from the experiment. 12 responses out of 24, which were classified as **canonical**, explained either, that charge/current built up over time, or that current oscillated, or that current repeatedly peaked and reached a low, or that it turned on and off repeatedly. While the last two responses did not use canonical language, they were classified as canonical, for the conceptual understanding, that current alternately fluctuated between maximum and minimum values, at regular intervals. The decrease in number of students who responded canonically before the experiment, from 26 to 12, was indicative of issues with interpreting the graphical output for current. 4 responses were designated **concrete-intuitive**, because although they referred concretely, to the match between the calculated and experimental values of time constant of current, there was no clear linkage established between the two. 5 other responses which referred directly to current, as wavelike or non-linear, reflected conceptual ambiguity with respect to what happened with current, and were also classified as **concrete-intuitive**.

Table 4.10 Experiment 4 Code Categories and Themes

Question	Theme	Code Category	Description	Count
PreQ1	Epistemic Construct	E.Expect_IBehinRC	Behavior of current when C is introduced in a circuit containing R	39
		Canonical		
	Concrete-Intuitive	K1.Iwilldec	Current will decrease	5
		K1.Wl stabilize when C charges	Current will stabilize when C gets charged	4
			Time to reach peak value will increase	10
		K1.TtoRchV/Rwline	Current/ circuit will switch on and off at intervals	5
		K1.I/CctWon-offatInt	Capacitor will store energy	2
		K1.CWlstoreE		
	Abstract-Intuitive	K2.changeV/I	Current/voltage will change	3
		K2.Vwilldec	Voltage will decrease	2
	Abstract-Intuitive	K3.Vwillinc	Voltage will increase	3
		K3.IwlInc	Current will increase	4
PostQ1	Epistemic Construct	E.Interpret_IBehinRC	How current behaved in RC circuit	28
	Concrete-Intuitive	K1.Ipeak-low/on-off,repeat	Current peaked and became low/turned on-off repeatedly	6
		K1.Q/Ibldsupovertime	Charge/current built up over time	2
		K1.Ifluc/oscatequalint/wtime	Current oscillates/fluctuates at equal time intervals	4
	Abstract-Intuitive	K2.TExptmatchesCalc	Experimental and calculated values of time constant matched	4
		K2.Inlin/behwavelike	Behavior of Current was not linear/behaved like a wave	5
	Abstract-Intuitive	K3.Iincwchar	Current increased with charging	3
PostQ2	Epistemic Construct	E.I,VBehinRC	Compare behaviors of voltage and current in RC circuits	21
	Concrete-Intuitive	K1.SameTConst	Had same time constant	3
		K1.V,Iinc-decprop/reg	Voltage and current increased and decreased proportionally/regularly	8
	Abstract-Intuitive	K2.V,IDirRel	Voltage and current had a direct relation	2
		K2.I,Vbeh/values-N/Ldiff	Values of I,V had little or no difference	6
		K2.V,Iinvrel	Voltage, current had an inverse relation	3

3 responses which incorrectly described current as having increased with charging, seemed to reflect either (i) the influence of a prior misbelief about the association between charging of capacitor and change in current, and/ (ii) difficulty in matching time periods on the graph which corresponded to charging of capacitor and increase of current. They were classified as **abstract-intuitive** for the misdirected abstraction caused by learners' own prior misconception.

PostQ2 asked students to compare the behaviors of voltage and current in RC circuits, based on the results of their experiment. Only 11 out of 21 responses were **canonical** – 3 students concluded that time constants remained same, and 8 concluded that the graphs increased and decreased proportionally, and/ regularly. While the descriptions of latter responses, lacked the use of canonically appropriate language, it was classified as canonical, for the conceptual understanding reflected.

Out of 8 responses classified as concrete-intuitive, 6 concluded vaguely, that the behavior and/ values of voltage and current were not very different. While they clearly based their conclusion on results of the experiment, it was unclear whether by 'same values' they were referring to time constants, and by same behavior, they were referring, for instance, to increase and decrease of both quantities at regular intervals of time. Hence, they were classified as '**concrete-intuitive**'. 2 responses which referred to a direct relation between V and I, and 3 responses which described the relation as inverse, were also classified concrete-intuitive, because the source of their knowledge, was evidently their experimental data, but the conclusions were more vague and intuitive, than canonical. It was unclear whether by 'direct relation' the students implied synchrony between output waveforms of V and I, and whether by 'indirect relation', they were implying that the increase and decrease of voltage and current were opposite, although synchronous.

To sum up, out of 39 students who responded to PreQ1, only 28 responded to PostQ1, and 21 to PostQ2, although they were merely asked to draw conclusions based on their observations. Thus, at least 25% of respondents to PreQ, chose not to respond, to an identical question (PostQ1) about the behavior of current in an RC circuit, after the experiment, possibly because it asked them to make inferences from graphical outputs. Second, only 50% of those who did respond, (14/28 for PostQ1, and 11/21 for PostQ2) framed the concept canonically. Most of the remaining 50% students, seemed to have left the lab with incoherent understandings of the working of RC circuits, and one of the main epistemic blocks which hindered coherent concept formation, was their inexperience and difficulty with reading and interpreting non-linear (or ill-defined) graphical outputs. These findings therefore, like those of previous experiments, again point towards a need to monitor and assess student thinking qualitatively, in-context.

Summary of Qualitative Findings

Qualitative analyses of learners' Twitter responses to one pre-question sent to them before each of the four experiments, and two post-questions, sent to them after each experiment, revealed four broad themes as follows: (i) Epistemic construct (E), or the physical concept which was constructed (described, justified, and/interpreted) by the students, using different knowledge elements; (ii) Canonical physics (K1), or the justifications, interpretations and descriptions which referred to canonically appropriate physical quantities, representations, and/relationships, in the given context; (iii) Concrete-intuitive physics (K2), or explanations which either referred to concrete embodiments (examples) of canonically appropriate explanations, or made somewhat intuitive, indirect references to canonical explanations, and whose origin could be traced back to concrete sources such as course readings; and (iv) Abstract-intuitive physics (K3), or explanations which could be interpreted as intuitive abstractions of canonically appropriate explanations, and which appeared to originate (source) in learners' misconceptions, and/misbeliefs, about the underlying physical mechanisms.

Experiment 1. The epistemic constructs framed by students for this experiment, were prediction of '(what happens when) voltage source (is held) in wet hands', before experiment, and interpretation of 'relationship between wetness and resistance of hand', after the experiment. While most students framed the concept canonically, as 'wet hands conduct better', or 'wetness decreases resistance of hands', before and after the experiment, many students also connected the concept to concrete-intuitive knowledge element, 'experience shock', before the experiment, possibly based on readings, which did include related content information, and/ their prior experiences from real life - warnings about electric hazard if you touch a live wire or voltage source, with or without wet hands. Few students were found to retain this notion, meaning these notions were not resolved through the inquiry activity. Thus, concrete-intuitive physics was the dominant theme before experiment, and canonical physics was the dominant theme after the experiment.

Experiment 2. The epistemic constructs framed by students for experiment 2 were, prediction of 'ways to change capacitance' of a parallel plate capacitor, before experiment, and description of 'ways capacitance was changed', and interpretation of 'relationship between capacitance, area, and distance' after the experiment. While most students framed the concept canonically as 'capacitance decreases with increase in distance', before, and after the experiment, less students made connections to the canonical relation between 'change in area', and 'change in capacitance', before, and after the experiment. Before the experiment, many students were found to predict physical relationships associated with change in capacitance, (abstract) intuitively, and/ vaguely in terms of physical quantities such as voltage, current, charge, weight, size, shape, orientation of plates, and almost half of these students retained these notions,

after the experiment. Thus, while canonical physics (relation between capacitance and distance), was the dominant theme for experiment 2, connections between ‘change in area’, and ‘change in capacitance’ remained either missing, or incorrectly connected in many learners’ thinking, even after the experiment. Also, abstract-intuitive physics was found to have significant presence in learners’ thinking, before and after the experiment.

Experiment 3. The epistemic constructs framed by students for experiment 3 were, explanation, classification, and justification of materials as ‘ohmic’ and ‘non-ohmic’. Most students framed the concept canonically before the experiment, as ohmic materials follow ohm’s law, exhibit linear relationship between voltage and current, and/have constant resistance; and non-ohmic materials do not follow ohm’s law, exhibit non-linear VI relationships and/ have non-constant resistance. However, after the experiment, approximately 30% of the students incorrectly identified light bulb as being ‘ohmic’, and 25% identified LED, incorrectly as ‘ohmic’, based on their incorrect interpretations of graph. Thus, while canonical physics was the dominant theme before the experiment, intuitive physics was also found to have significant presence in learners’ thinking, after the experiment, and learners’ incorrect abstractions could be traced back to their incorrect interpretations of linearity of graphs.

Experiment 4. The epistemic constructs framed by students for experiment 4 were, prediction and description of ‘behavior of current in RC circuits’, and comparison of ‘behaviors of voltage and current in RC circuits’. Most students framed the concept canonically before the experiment, as current will decrease, time to reach maximum value will increase, circuit/current will stabilize when capacitor charges, at least one third of the students were found to have intuitive notions about the behavior of current in RC circuits. After the experiment, 50% of total responses reflected canonical physics, and 50% reflected intuitive physics. Thus, while the dominant theme for experiment 4, was canonical physics, before the experiment, 33% of respondents, did reflect dominance of ‘intuitive physics’ (split equally between concrete and abstract) in their thinking. After the experiment, however, 50% of responses were characterized as canonical physics, and the other 50% were characterized as intuitive physics, with concrete-intuitive theme being more dominant than abstract-intuitive.

Coded qualitative data, generated by qualitative analyses presented in this section, was formatted, per the requirements of ENA WebTool, and exported to it for further analyses. ENA Web Tool transformed the coded data into adjacency matrices and network graphs, which were analyzed using Network Analysis by the ENA WebTool. The findings of Network analyses, which include qualitative as well as quantitative techniques, conducted on the transformed data, will be presented in the next section.

Network Analysis Findings: (QUAL+QUAN)

While the qualitative findings as reported in the previous section, do reveal substantive information about broad trends in students' thinking before and after the experiment, qualitative techniques alone, are not adequate to conduct further investigations by groups, for instance, which are critical for assessing learner differences, especially since equity of access to 'the physics' is an important issue in the domain of physics. Also, as previously discussed (Chapters 1 and 2), looking at 'co-occurrences', or 'adjacency', or 'connections' rather than mere occurrences of specific elements of knowledge, skills and/beliefs, will reveal richer information about epistemological factors which could be hindering learners' access to canonical knowledge. Network analysis offers a powerful analytical tool to visualize, quantize, and statistically analyze such information, without losing its qualitative details, and richness, as is illustrated by the results of network analysis discussed in this section. These findings illustrate and lend support to the argument that network analysis can enhance the power of qualitative techniques, by essentially facilitating qualitative and quantitative analyses of large sample data.

The qualitative themes and categories identified in each 'Twitter response', and the text of the response itself, were catalogued for each learner by experiment and question, with the help of spreadsheet, as discussed in the previous section. This spreadsheet data was then merged with the 'metadata', containing information about each learner's quantitative scores and group membership (high, low, medium, determined through quantitative analyses presented in Section 1 of this chapter) on five quantitative instruments – EBAPS, Pre-lab home-work quiz (attempt1), LabQuiz4, Lab4, and Test1. Merged data for each experiment and question was then formatted as per the requirements of ENA Web Tool (using identifiers for metadata, raw data, and code columns), on a separate Spreadsheet file, saved in the '.csv' file format, and uploaded onto the ENA WebTool. Four combined files were also created and uploaded for each experiment, by merging together all data, to enable pre-post network analyses.

For each csv file, after I assigned the 'unit', 'conversation', and 'codes' to be modeled as network graphs and adjacency matrices, the ENA WebTool generated a separate ENA set (rData). Two types of analyses were then conducted: (i) Network graphs of each ENA set, plotted by the ENA WebTool, were analyzed qualitatively and quantitatively (ii) The rData file of each ENA set created by the ENA WebTool, was downloaded to my computer, saved, and then quantitatively analyzed using the open source statistical computing software, RStudio. The network graphs were analyzed qualitatively and quantitatively (t-tests). In contrast to qualitative analyses presented in previous section, network graph analyses yielded quantizable, yet qualitatively rich information about learners' thinking, potentially useful for formative classroom assessments, to inform instruction design.

In this section, I will present the findings of network graph analyses, by Experiment (Qual + Quan), to answer RQ2: (a) What mean patterns of connections do the Network Graphs reveal in learners'

thinking before and after each of the four inquiry experiments (activities)? (b) Do the pre-and post-activity patterns differ significantly by (i) Section of course and (ii) Levels of (a) Epistemic beliefs Score (b) Lab Quiz Score (c) Lab Score (d) Homework Quiz Score (Attempt1) and (e) Test 1 Score? (QUAL+QUAN).

Experiment 1 (The Resistance of Your Body)

This sub-section will present Network representations of connections (edge of network) identified in learners' thinking, by coding co-occurrences of codes (nodes of network), and report the findings of qualitative and quantitative (between group t-tests) analyses of these networks, for responses to PreQ and PostQ, of Experiment 1, which asked learners to explain and interpret the relationship between wetness of hand, and its resistance. Analyses of pre-post differences in mean epistemic networks, individual student epistemic networks, and comparison of mean networks by Section of course, and by Low medium, and High score groups, for five quantitative instruments, will be presented. However, the network graphs and tables of t-test results, for statistical significance of differences in mean epistemic networks of students grouped by their scores, on EBAPS, A3 Home-work Quiz, Lab Quiz, Labs, and Test1, will be appended in Appendix 5, 6, 7, and 8, respectively, for experiments 1, 2, 3, and 4.

(a) Pre-Post Mean Epistemic Networks of all Learners

In Figure 4.1 below, each centroid (red and blue solid dots) on the left network graph, corresponds to the mean location of the weighted epistemic network of one student in the network projection space, for either PreQ (red) or Post Q (blue), along two dimensions – 1(X) and 2 (Y) which present maximum variance for network visualization. The solid red and blue squares are the positions of mean networks for PreQ and Post Q respectively, and the rectangles around them indicate confidence intervals. The network graph on right shows the mean epistemic networks of all students for PreQ (red) and PostQ(blue). The thickness of each edge (network connection) is proportional to strength of association (connection) between the two nodes (knowledge elements) connected by it, which in turn depends on the number of times the nodes co-occurred in learners' responses. The two graphs show that there were significant differences in cognitive connections made by students before and after the Experiment, along Dimension 1. The t-test results for the two samples (student networks before and after experiment), are shown in Table 4.11 below.

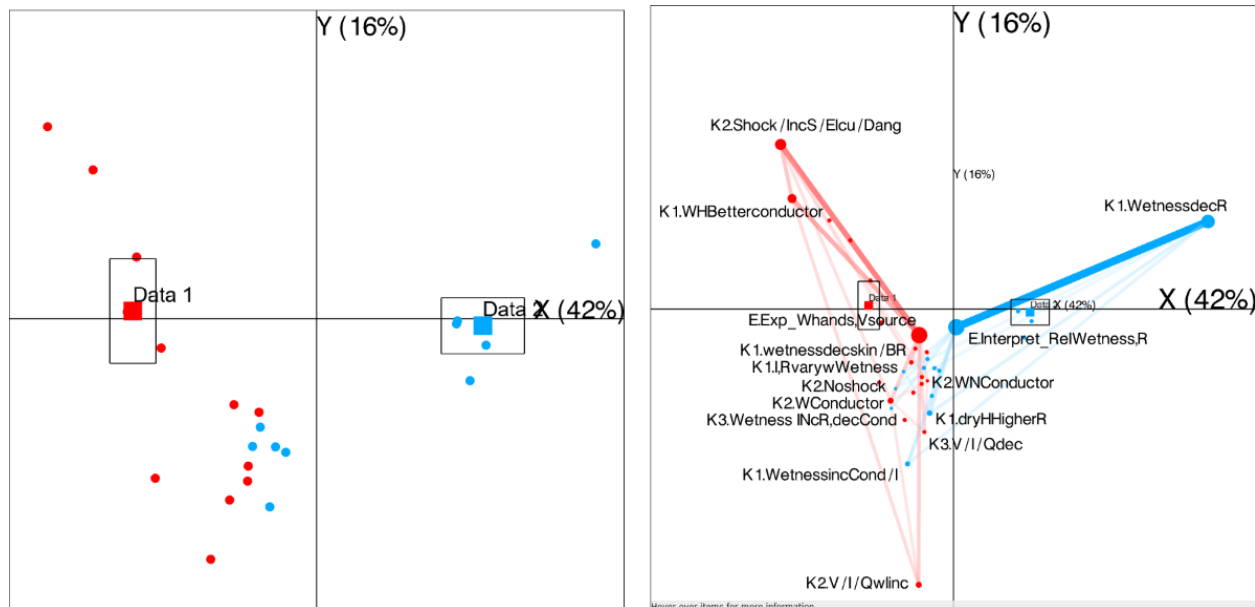


Figure 4.1. Centroids of cognitive networks of students with means and confidence intervals (left graph) and Mean Networks (right graph) for PreQ(red) and PostQ(blue).

Table 4.11 T-test for Significance of Differences (PreQ and Post Q Connections)

	Dimension 1(42)	Dimension 2 (38)
Data 2 (Post-Q networks)	0.404	-0.017
Data 1(Pre-Q networks)	-0.447	0.019
T statistic	14.87	-0.5
P value	< 0.001	0.619
Cohen's d	3.276	-0.116

At an alpha level of .05, P-value is less than .001, and the effect size (Cohen's-d) is greater than 0.8, which means that the two samples - PreQ and Post Q mean network models, have statistically significant differences. The right graph shows that on an average, before the experiment, students made strongest connections between the 'epistemic construct' and concrete-intuitive knowledge element (KE) 'shock', and less strong connections with canonical KE – 'wet hands better conductor'. After the experiment, strongest connections were made to canonical KE – 'wetness decreases resistance'. Thus, the experiment was instrumental in shifting learners' epistemic frame, towards thinking canonically in terms of resistance of hand. It should be noted that the Post Q explicitly directed their thinking towards the results of experiment, while the PreQ was broad and sought to test their understanding from the readings, and this difference in 'specificity' could have contributed to the observed differences in network connections.

(b) Pre-Post Individual Student Epistemic Network Examples

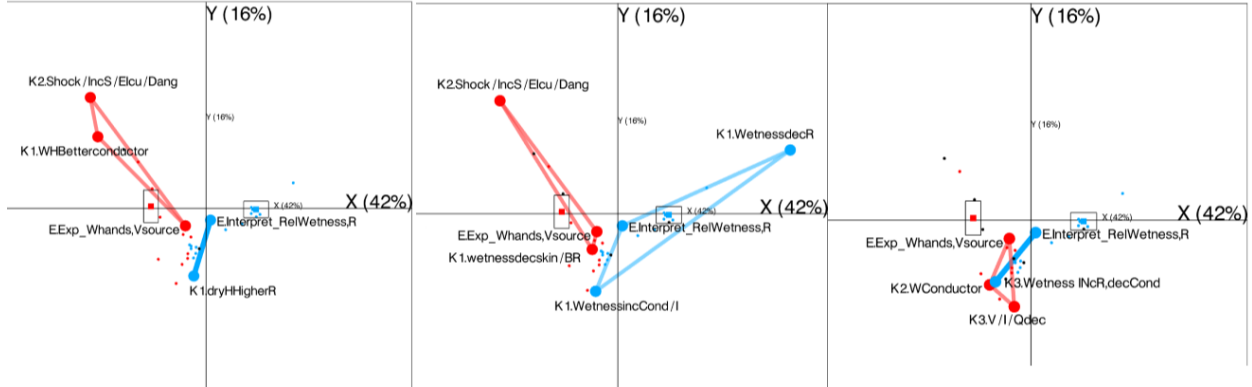


Figure 4.2. Epistemic Networks of student F23 (left graph), student F32 (middle graph) and student F31 (right graph) for PreQ (red) and PostQ (blue).

The epistemic network of F23 before experiment consisted of connections to both canonical (K1) and concrete-intuitive (K2) knowledge elements, while after experiment, she connected the epistemic construct, strongly, to canonical KE ‘dry hands have higher resistance’ (Figure 4.2). F32’s network also changed similarly but had more extensive connections to canonical physics after experiment. F31, connected both to concrete-intuitive knowledge element (KE), ‘water is a conductor’, and to abstract-intuitive KE, ‘voltage/current/charge will decrease’, before experiment; but after the experiment, she connected the epistemic construct to abstract-intuitive KE - ‘wetness increases R, decreases conductance’. To understand her reasoning, I looked at her group’s Lab Report for Experiment 1. I found that they did not have complete records of resistance of ‘dry’ and ‘wet’ hands, and their prediction of current appeared to be based on a value of resistance 5.2 megaohms, which could correspond to dry hands. It is possible that they were unable to record observations for wet hands, simply because they did not scale down the setting on multimeter from megaohms (TA comments on Lab Report pointed this out). Hence, she lost the opportunity to shift her abstract-intuition towards concrete-intuitive, or canonical KEs, through the inquiry. This ‘issue’ could have been caught or captured in real time, using the ENA approach proposed.

(c) PreQ and PostQ Mean and Individual Student Network Models

PreQ. Left graph in Figures 4.3 represents the positions (centroids) of mean epistemic networks of individual students for PreQ (small red solid dots) in the network projection space, and the small and big squares represent the mean network and its confidence interval, for all students. The number of ‘dots’ is smaller than the number of students, because more than one student’s epistemic networks occupy the same position in the network projection space, and their networks resemble each other very closely. The right graph in figure 4.3 represents the expanded epistemic network of mean centroid (small solid square)

from the left graph. It gives an aggregated view of the structure of network connections of all students before the experiment. The results were previously discussed with PreQ above (figure 4.1, right graph, red network).

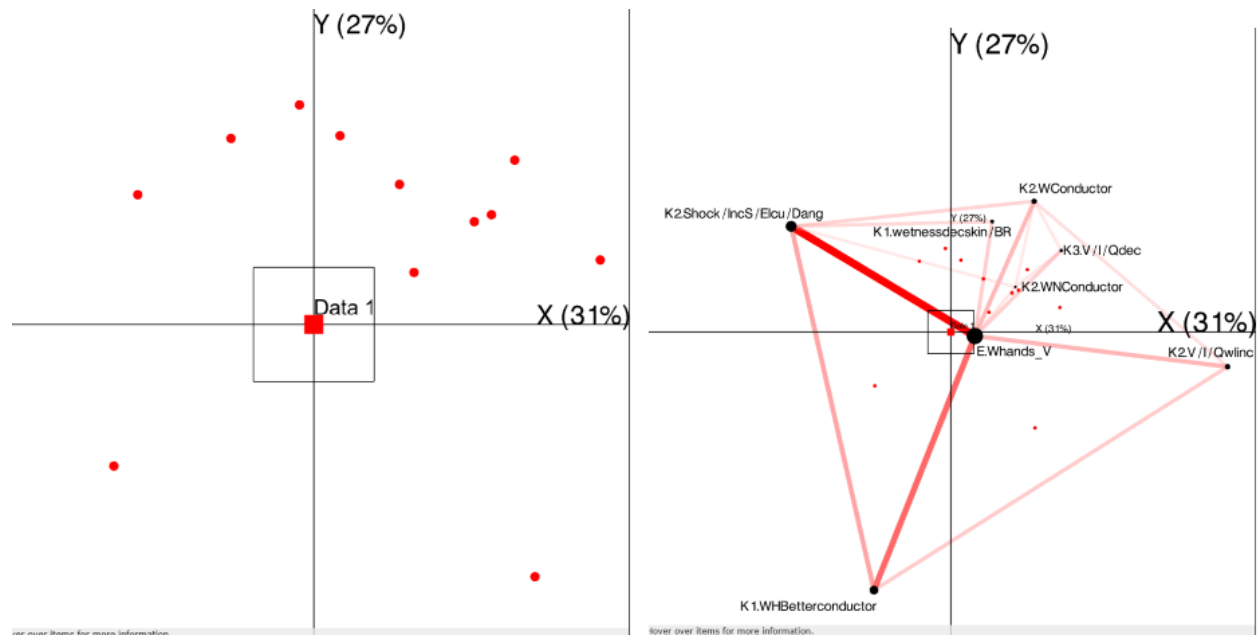


Figure 4.3. Mean and individual student centroids (left graph) and mean epistemic network (right graph) for PreQ.

Graphs in figure 4.4 below, represent the individual networks of students by the position of their centroid on the left graph (LG) of figure 4.3 above. Most students' centroids were in Quadrant 1(+X, +Y) of the graph, (figure 4.3 LG), 3 centroids were located in Quadrant 2(-X, +Y), and one each in Quadrants 3 (-X, -Y) and 4 (+X, -Y). By cross referencing figure 4.3LG, with Figure 4.4 LG below, it was found that students with centroids in Quadrant 1 of figure 4.3LG, made most connections with concrete and abstract-intuitive KEs (K2 and K3), resulting in a score of '0' for K1-centrality (centrality of K1 KEs in thinking), and for canonical and transitional connections scores (CCS and TCS), calculated in the next (quantitative results) section.

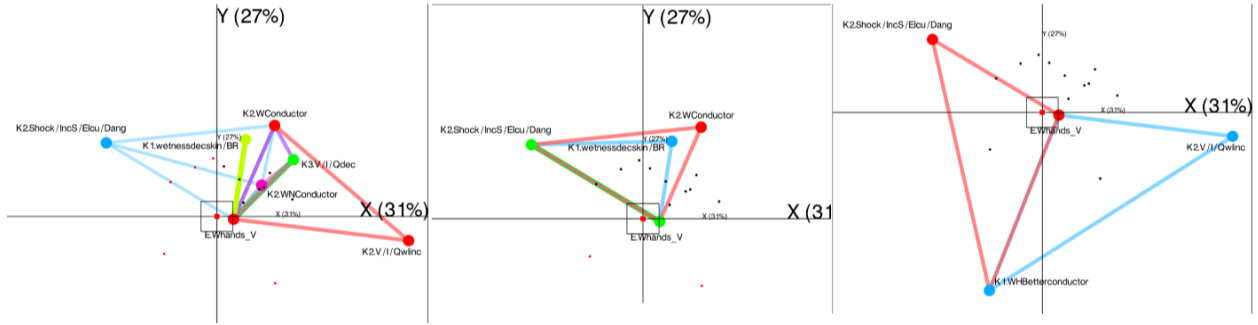


Figure 4.4 . Individual student models with centroids in Quadrant1(left graph), Quadrant 2 (center graph), and Quadrants 3 and 4(right graph) for PreQ.

Different colors represent networks of individual centroids (of one or more students). Students with centroids in Q2 of graph (Figure 4.3LG), connected strongly with concrete-intuitive, and less strongly with canonical KEs, (Figure 4.4 CG) resulting in a score of ‘0’ for abstract-intuitive connections (I2CS), and K3-centrality discussed in the next section. Student centroids in Q3 and Q4 of Figure 4.3LG, demonstrated similar networks with co-occurrence of K1 and K2, and absence of abstract-intuitive KEs before the experiment.

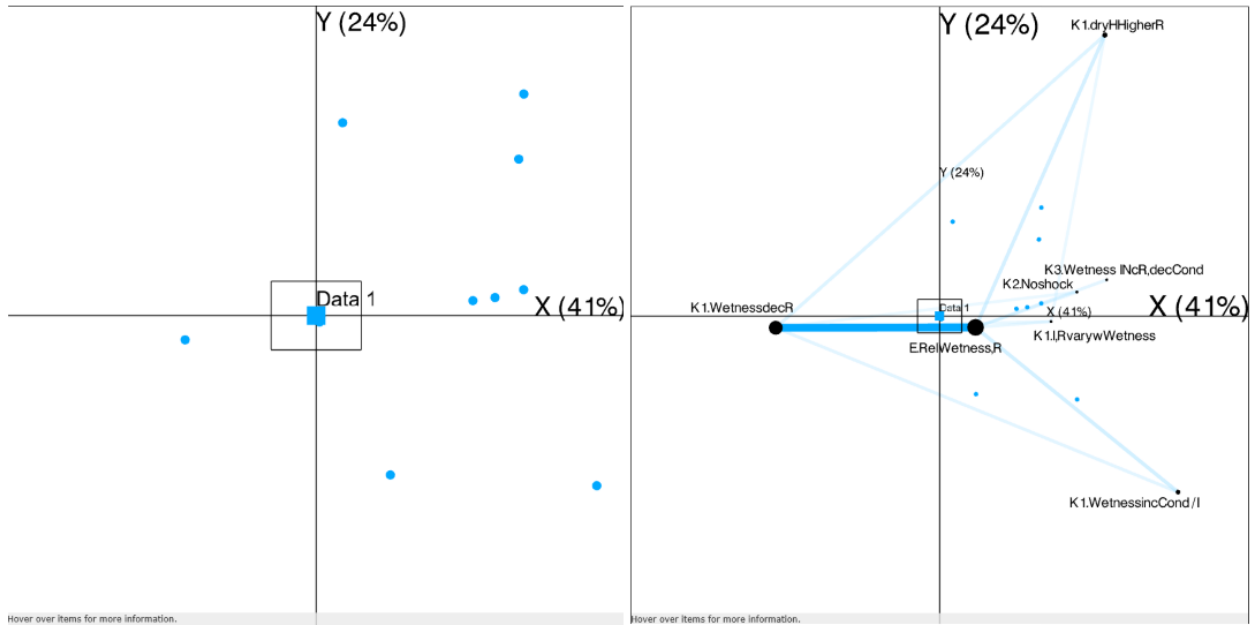


Figure 4.5. Mean and individual student centroids (left graph) and mean network model (right graph) for PostQ.

PostQ. Similar to PreQ above, Figure 4.5LG represents the centroids and mean centroid (solid blue square at center) of student networks modeled from their responses to PostQ, and Figure 4.5RG represents the mean epistemic network of all students for PostQ, which showed that students made the most connections to canonical KE, ‘wetness decreases hand resistance’ (thick blue line in Figure 4.5RG) and weak, or significantly less connections to concrete-intuitive KE, ‘experienced no shock’, and to abstract-intuitive KE, ‘wetness increases resistance or decreases conductance of hand’. These results were also obtained from the Qualitative analyses of previous section, with a major difference, that those analyses did not consider co-occurrences, and required looking at ‘raw data’ to determine the presence or absence of other KEs. Network graphs, on the other hand, present a rich visual presentation, amenable to qualitative interpretation, as well as quantitative comparisons (done later in this section).

Cross -referencing left graph of Figure 4.5 above, with graphs in Figure 4.6 below, it is seen that individual epistemic networks of centroids in Quadrant 1 of Figure 4.5LG, are represented by left graph of Figure 4.6, and individual epistemic networks of centroids in Quadrants 3 and 4 of figure 4.5LG, are represented by right graph of Figure 4.6. Networks in left graph, and therefore, students with centroids in Q1 of 4.5LG, were found to have made strong connections of the epistemic construct ‘relation between wetness and resistance of hand’, to canonical, concrete-intuitive, and abstract-intuitive KEs, while those in right graph, connected the concept to only canonical KEs, meaning these students (with centroids in Q3 and Q4 of Figure 4.5 LG), demonstrated a higher level of sophistication of thinking.

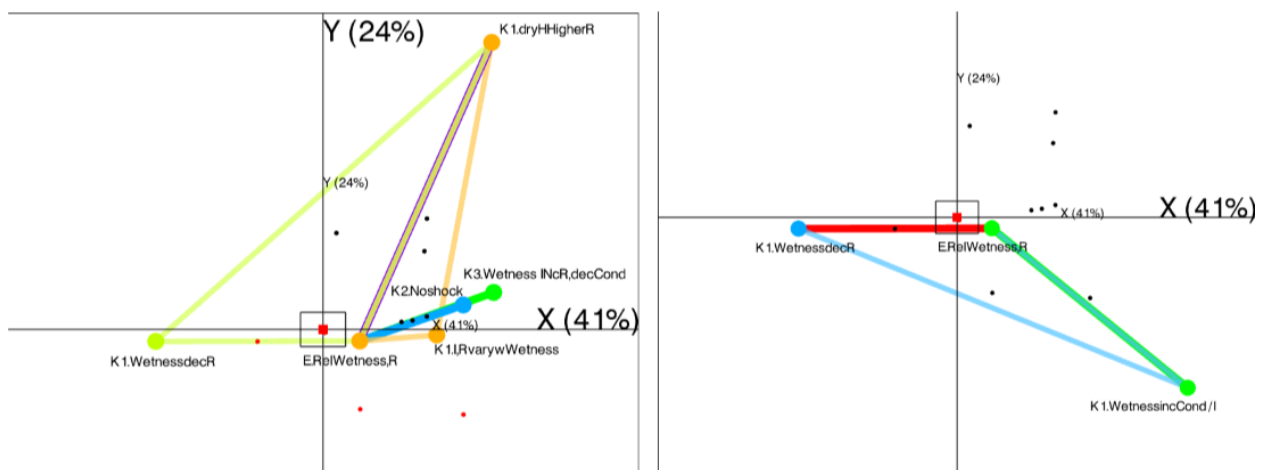


Figure 4.6. Individual student models with centroids in Quadrant 1 of 4.5LG (left graph), and Quadrants 3 and 4 of 4.5LG (right graph) for PostQ.

These same results can also be inferred by cross-referencing the left and right graphs of (i) Figure 4.5 for PostQ, and Figure 4.3 for PreQ, or (ii) Figure 4.1 for both PreQ and Post Q. For instance, Q1 of right graph of 4.3, shows stronger connections to K2 and K3 nodes, than to K1. The centroids of students who made those connections can be found in the corresponding quadrant, Q1, of left graph of 4.3. Similarly, in Figure 4.5 above, two K1 nodes with somewhat strong connections to the epistemic construct, are located in Q 3 and Q 4 of right graph each, and their corresponding student centroids (mean positions of their networks in the projection space) can be found in the corresponding Quadrants 3 and 4 of left graph. Students' names or IDs can be found on the bottom pane of the 'active graph' in the ENA WebTool, when the cursor (mouse of computer) hovers over the centroid, and the students' 'response' or raw data opens in another window, on clicking a specific edge or connection. Thus, the ENA Web Tool interface offers a dynamic view of individual, and/ aggregated data, for each ENA set created.

(d) Pre-post Network Model Comparison by Section of Course

Figure 4.7 below presents the centroid and network graphs of Section 1(red) and Section 3 (blue) for PreQ (upper left, center, and right graphs -ULG, UCG, URG), and for post Q (lower left, center, and right graphs - LLG, LCG, LRG).

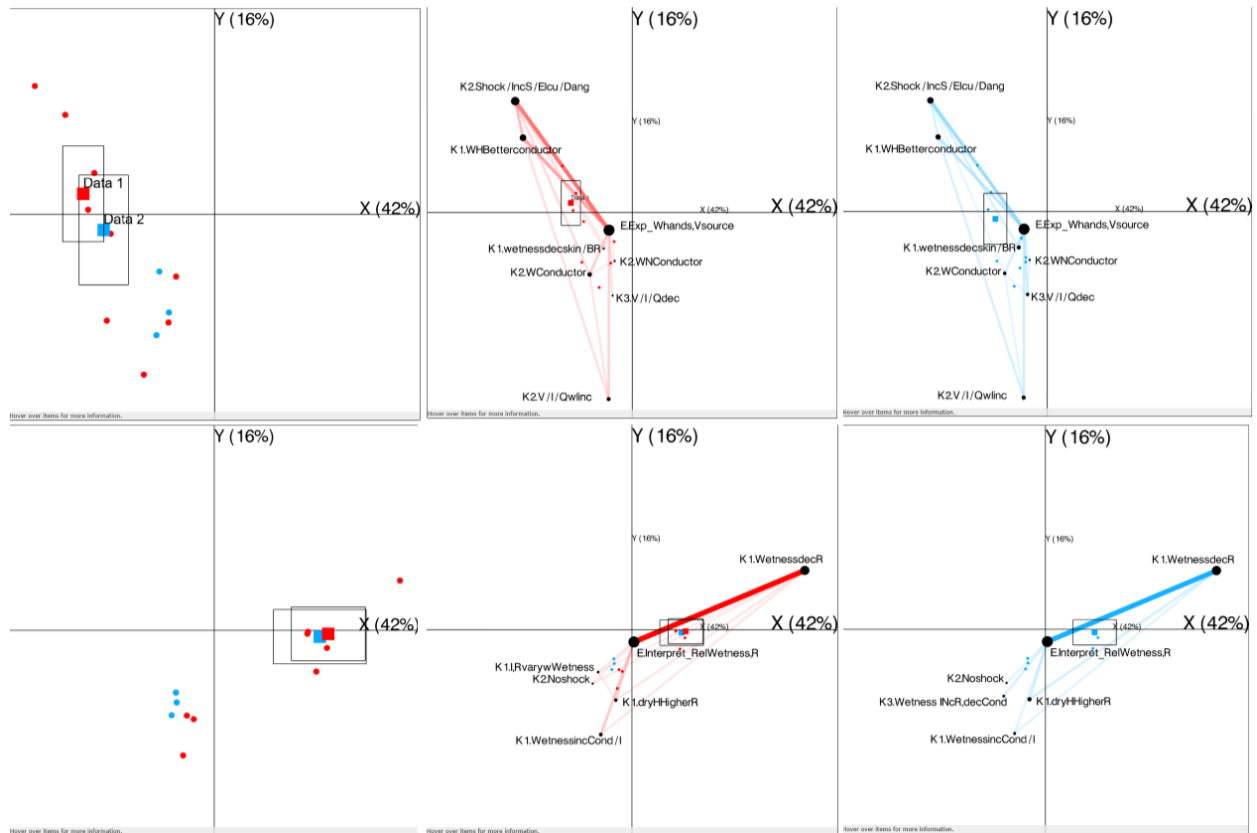


Figure 4.7. PreQ centroids (left upper graph) and Mean networks by Section - Section1(red), Section 3(blue) (center and right upper graphs); PostQ centroids (left lower graph) and Mean networks by Section – Section1(red), Section 3(blue) (center and right lower graph).

As can be seen from the results of t-test (Table 4.12 below), there were no statistically significant differences between the two sections, either before or after the experiment, because p-values are greater than .05 for both PreQ and Post Q, although a medium effect size for PreQ (Cohen's d greater than 0.3), indicating a slightly higher mean for Section 3 along dimension 1. The confidence interval overlap, which can be seen in the upper (PreQ) and lower (Post Q) left graphs, along both dimensions (x and Y axes), also points towards the same result. One qualitative difference which stands out, is that after the experiment, one or more students from Section 3, retained unresolved abstract-intuitive KE, 'wetness increases R, decreases conductance', and another was that, canonical KEs found in the mean network of Section 1, were more diverse, although, before the experiment, both sections shared almost identical network connections.

Table 4.12 T-test for Significance of Differences in Epistemic Networks by Section of Course – Section1, and Section 3 (PreQ and Post Q)

	PreQ		PostQ	
	Dimension 1(16)	Dimension 2 (22)	Dimension1(25)	Dimension 2(17)
Section 3	-0.404	-0.057	0.386	-0.023
Section 1	-0.478	0.074	0.417	-0.013
T statistic	1.34	-1.046	0.301	0.159
P value	.19	.303	.765	.875
Cohen's d	0.452	-0.35	0.097	0.049

(e) Epistemic Network Comparison of Groups Differentiated by Quantitative Instruments

The following analyses were conducted to explore whether and how the statistical t-test feature of ENA Web Tool could be used to evaluate and inform equity of instructional measures, by enabling various types of group comparisons. One objective of these analyses, was to explore for any correlations which could be identified, between learners' performance as assessed by existing instruments, and the sophistication of their thinking, as captured by ENA-based assessment of Twitter responses. A second objective, was to test the reliability of ENA-based assessment to measure learner differences, by using the assessment technique repeatedly across four different experiments, to measure for significant differences between learners, grouped by their performance on other assessments. Discussion will therefore, be limited to summarizing the main descriptive results (of t-tests), and qualitative discussion of note-worthy differences, and/similarities. Table of t-test results, and network graphs are appended in Appendix 5.

EBAPS. The average EBAPS score for this sample was 1.63, with a high of 1.8. In comparison to an Expert score of 4, the participants could be said to have novice-like beliefs about physics. The premise of this study was that learners taking the course presumably had contextual epistemologies, which resulted in their performance being locally, rather than globally coherent, and therefore, their epistemic frames needed to be monitored and attended to. The results of EBAPS confirmed that prior to taking the course, and conducting the activities, learners' epistemological beliefs about physics were indeed novice-like.

Hence, it was expected that there would be no significant differences between mean epistemic network models, by EBAPS scores, if the learners had novice-like beliefs, and therefore, contextual epistemologies. The results of t-test for significant differences (Table 5.1, Appendix 5) confirmed, that there were no statistically significant differences between the network models of students, grouped by their EBAPS scores, before and after the experiment. Figure 5.1 (Appendix 5) presents the mean network graphs of Students grouped by their EBAPS scores – Low (overall score less than or equal to 1.5), Medium (overall score between 1.5 and 1.7) and High (overall score greater than or equal to 1.7). The

upper three graphs represent mean networks of the three groups for PreQ, and lower three represent mean networks of the three groups of students, for PostQ.

A3 Home Work Quiz (Attempt1). In this course, multiple attempts are permitted on the Homework quizzes, which are taken online by students, ahead of the studio sessions. They consist of multiple choice conceptual problems, graded automatically by the course management system, and the purpose of providing multiple attempts, could be, that students review the readings and get their concepts clarified. Since the final quiz score was unlikely therefore, to differentiate between students' knowledge levels or prior beliefs (which could be affecting their performance on the Quiz), I decided to group students by their scores on the first attempt, which were more likely to capture these effects. The null hypothesis for a t-test, by default, is that there are no significant differences between group means. The results of t-test (Table 5.2, Appendix 5) found the existence of statistically significant differences before the experiment, at an alpha-level of .05, between Low-score (less than or equal to 5) and High-score (equal to 8) groups, and the corresponding network graphs (Figure 5.2, Upper left and right graphs, Appendix 5) revealed structural differences in networks. The graphs show that high scorers made connections to abstract-intuitive KE, 'V/I/Q decrease', along with canonical KEs, in contrast to low scorers, who framed the question in terms of canonical and concrete-intuitive KEs, reflecting a more sophisticated level of thinking, than the high scorers.

The finding, that although the three groups connected to canonical KEs, high scorers also connected to concrete-intuitive KE, 'water not conductor' and abstract-intuitive KE, 'voltage/current/charge will decrease', concurrently, supports the argument of this study, that quantitative assessments, which focus merely on determining presence of canonical KEs, fail to uncover and account for 'connection issues in learners' thinking', which therefore, remain unaddressed by the TA/instructor. Hence the need for qualitative assessments based on ENA. The statistically significant differences between low and medium (scores equal to 6 and 7) groups, after the experiment, also were similar to low-high differences before experiment and illustrated 'issues in thinking' of medium scorers as compared to low scorers. It could also be, that the strategy of allowing multiple attempts on the homework quiz, was effective, in the sense that it incentivized the low performers to review the content more thoroughly than they did before, and therefore, their epistemic frames, reflected no abstract-intuitive KEs, before the experiment, and reflected only canonical KEs after the experiment.

It should be noted, however, that the low scorers for this quiz, were significantly less in number (4) as compared to the high (20) and medium (16) performers, and the t-test results by themselves, cannot therefore, be used as indicative of differences. They did however, draw attention towards the networks, and reveal instructionally useful information.

Lab Quiz. The Lab quizzes are pen and paper, closed-book tests, consisting of problems and/questions pertaining to the readings for the labs, and are graded by the TAs. At an alpha-level of .05, statistically, there were no significant differences found, using the t-test for significance of null hypothesis (no differences), between student groups whose Lab 4 Quiz scores were sorted as low (less than or equal to 7), medium (7.5 to 8.5), and high (equal to 9 and 10) (Table 5.3, Appendix 5). The network graphs of low Lab Quiz scorers for PreQ, however, do show the connections to abstract-intuitive KE, ‘voltage/current/charge will decrease’, unlike the A3Quiz low score groups (Figure 5.3, Appendix 5). This finding suggests that ENA based assessments could also present as a tool for comparative evaluation of the ability of other quantitative assessments to mirror, and/capture sophistication of thinking. [It should be noted here, that the nonparametric correlation analysis using Spearman’s rho correlation coefficient, for all existing assessment instruments, revealed that there was small correlation between Lab Quiz and Lab scores (0.380), between Lab 4 and Test 1 scores (0.320), and high correlation between A3 Quiz (attempt 1), and A3Quiz scores (0.635).]

Lab Scores. T-test results of Table 5.4 (Appendix 5) indicated no significant differences between mean epistemic networks of learners grouped by their Lab scores - low (less than or equal to 85), medium (86 to 96), and high (greater than or equal to 97), since the p-values are all greater than 0.05, at an alpha-level of 0.05. Figure 5.4 (Appendix 5) shows that the network models of students in all three groups were identical before the experiment, but the low Lab score group shows no connections to abstract-intuitive KEs after the experiment, compared to the medium and high score groups. These findings are like those found with A3 Quiz, although unlike A3Quiz, all three groups of Lab 4 scores, have approximately equal number of students (10 to 13 each).

Test1. As revealed by the results of the t-test (Table 5.5, Appendix 5), no significant differences were found in the mean network models of students grouped by Test 1 scores, as ‘Low’ (less than or equal to 9), ‘Medium’ (10 to 12), and ‘High’ (greater than or equal to 13), for both PreQ and PostQ. All p-values were greater than 0.05, at an alpha-level of 0.05. Despite the small correlation between Test 1 and Lab 4 scores (0.320), the network structures of low, medium and high score groups of the two assessments, were found to be different for PostQ. While the network structures did not differentiate by Lab4 scores, and all 3 groups had identical network structures for PreQ, with K1-K2-K3 connections, there was evidence of structural differentiation by groups, in the case of Test1 scores (Figure 5.5, Appendix 5), for both PreQ, and PostQ. For PreQ, low and high Test 1 scorers showed K1-K2-K3 connections, but networks of high were more diverse (more K1). For Post Q, low Test 1 group showed predominantly K1-K3 connections, medium group had K1-K2, and high scores group had K1-K2-K3 scores.

Experiment 2 (Measuring capacitance)

PreQ for this experiment, asked students to explain ‘how the capacitance of a parallel plate capacitor could be changed’; Post Q1, asked them to describe ‘how they changed the capacitance in the experiment’ and PostQ2, asked them to interpret ‘the relation between capacitance, distance, and area of a parallel plate capacitor’. In this sub-section, I will present the findings of network graph analysis, on the same lines, as was done for Experiment 1 above. As before, I will draw upon these findings, to illustrate the benefits of using ENA as an interpretive tool for formative assessments in introductory university physics courses.

(a) Pre-Post Mean Epistemic Network Model Comparison

Although the target of all three questions was learners’ thinking about the concept, ‘change in capacitance’, PostQ2 was worded differently in comparison to PreQ and PostQ1, and it was expected that students’ responses to PostQ2, would focus on the relationships between physical quantities, unlike PostQ1 and PreQ, whose responses were expected to focus on identification of the ‘quantities themselves’. So, only the results of t-test for significance of differences between student networks for PreQ and Post Q1, are presented below (Table 4.13). At an alpha-level of 0.05, it was found that the two samples – student networks for PreQ, and student networks for PostQ1, differed significantly, along both dimensions, since the p-values were less than 0.05, although the effect size, or the size of the difference is greater along Dimension 2 (Cohen’s $d=1.118$), than along Dimension 1.

Table 4.13 T-test for Significance of Differences in Epistemic Networks between PreQ and Post Q1

	Dim 1	Dim 2
Mean PreQ	-0.269	0.008
Mean PostQ1	-0.045	-0.319
T statistic	-2.563	4.97
P value	0.013	<.001
Cohen’s d	-0.572	1.118

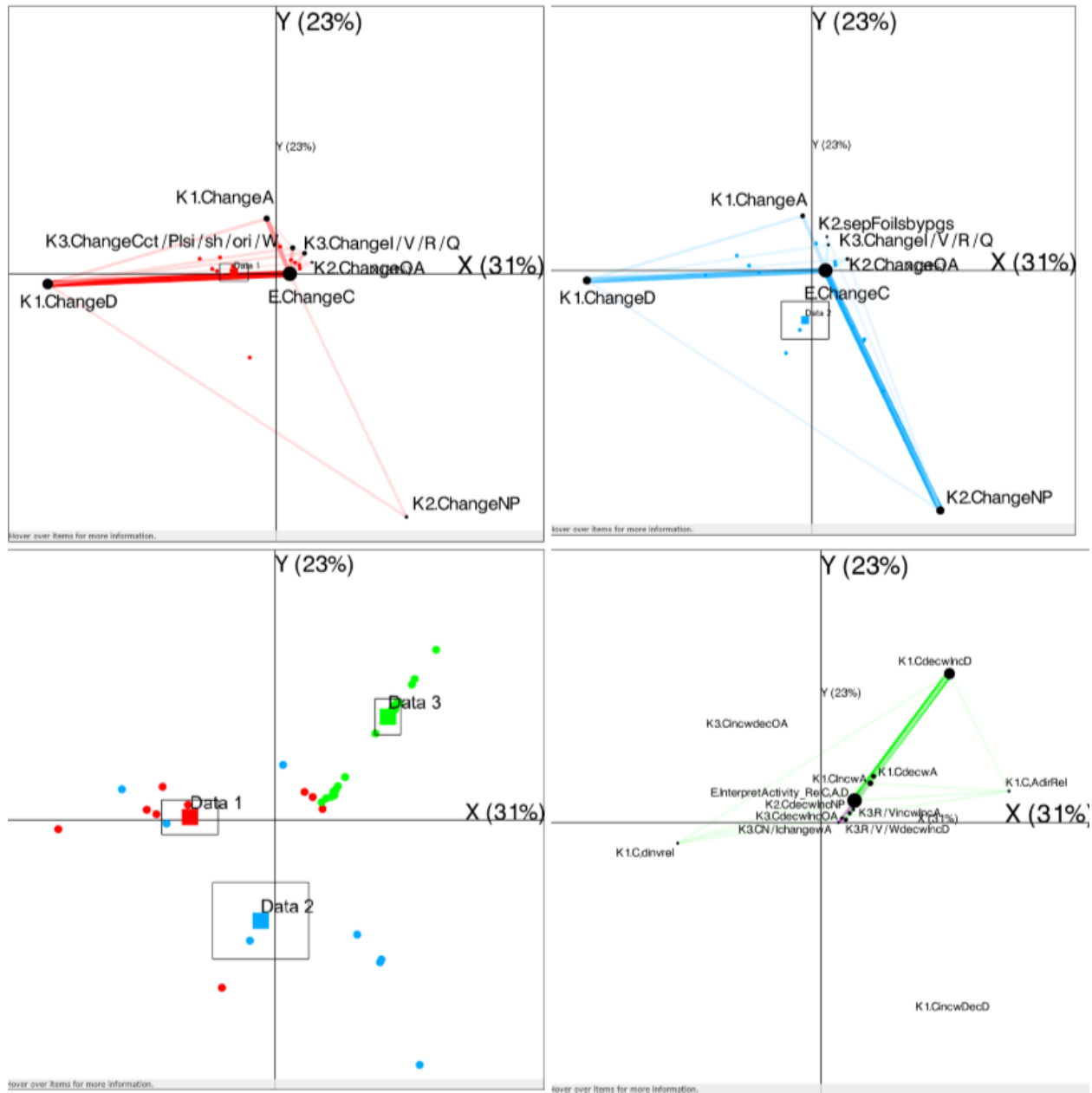


Figure 4.8. Mean networks for PreQ (Upper Left in red), PostQ1 (Upper Right in blue), PostQ2 (Lower Right in green), and centroids for PreQ (Data1), PostQ1(Data2), and PostQ2(Data 3) (Lower Left graph).

This was also confirmed by the confidence intervals of centroid positions for PreQ (red) and PostQ1 (blue), on the lower left graph in Figure 4.8 above, which are more separated along Y (Dimension 2), than along X (Dimension 1). The network graphs of Figure 4.8, present a comprehensive view (both qualitative and quantitative) of the structure of students' connections, which captures in one figure, the substance of qualitative findings, presented in the previous section, for Experiment 2. At the same time, network analyses also hold potential for a wide range of quantitative group comparisons and/ individual analyses, while retaining detailed qualitative information in-tact.

Comparison of upper 2 graphs of Figure 4.8 reveals that before the experiment, students made strong connections to the canonical KE, 'change distance' and less strong connections to the canonical KE 'change area' (Upper Left Graph), and, that the experiment was not effective in shifting their thinking to include the KE 'change area' (Upper Right Graph). However, students, not surprisingly, began thinking in terms of 'change number of pages', along with 'change distance', since they did change distance by changing the number of pages in the experiment, but did not similarly connect 'change overlap area' to 'change in area'. The lower right graph (PostQ2), also supports this finding, since it reflects strongest connections to 'capacitance decreases with increase in distance', while mixed types of connections, more intuitive, than canonical, were found with respect to relationship between capacitance and area. Thus, students will benefit from more direct 'intervention' by the TAs, to draw their attention towards the 'change in area' which took place, when they changed the overlap area.

(b) Pre-Post Individual Student Epistemic Network Models



Figure 4.9. Epistemic networks of students F33 (left graph), F14 (center graph), and F8 (right graph) for PreQ (red), PostQ1 (blue), PostQ2 (green).

Network graphs offer a dynamic view of shifts in an individual learner's thinking, before, during, and after an instructional activity. Figure 4.9 above, presents individual network graphs of three students, with different types of qualitative changes in their thinking, before and after the Experiment 2. Each graph shows the connections made for PreQ, PostQ1, and PostQ2, in red, blue, and green colors respectively. F33's epistemic frame shifted from canonical before experiment, to concrete-intuitive for PostQ1, and abstract-intuitive for PostQ2. The students' responses for each experiment were as follows:

PreQ - *"We could change capacitance of parallel plate condensers by decreasing/ increasing the space that is between the condensers. This will increase or decrease the effect they have on one another"*

PostQ1- *"We changed the capacitance in the experiment by increasing and decreasing the number of pages that were between the capacitors. The more pages between the capacitors, the less the resistance was and the less pages in between them, the more resistance there was in the system"*

PostQ2 - *"We learned that the greater the distance between the plates, the less effect they had on each other, and the less resistance there was. The area did not seem to have much of an effect on the capacitance. However, we believe that the larger the area of the plate, the greater the capacitance it should have."*

The epistemic frame of F14 shifted from abstract-intuitive, 'change circuit orientation', to concrete -intuitive, 'change separation levels between foils', for Post Q1 to what I will classify later in the section, as 'transitional connections'- combination of canonical KE, 'capacitance decreases with increase in distance', and abstract-intuitive KE, 'capacitance decreases with increase in overlap area'. Student's responses were as follows:

PreQ – *"by changing the capacitor to a series plate instead of a parallel"*

PostQ1 – *"we used different levels of separation between the two sheets of foils to change the capacitance"*

PostQ2 – *"The distance of the plates decreased the capacitance and so did the area of the sheets"*

In the case of student F8, the shift was from transitional connections – combination of K1, K2 and K3, to concrete-intuitive for PostQ1, and canonical for PostQ2. Students' responses were:

PreQ – *"Flipping a switch on, fully charging the capacitance of the parallel plate. As well, the plate spacing of the capacitors and the area of the capacitors... all other factors being equal, greater plate area gives greater capacitance and vice versa"*

PostQ1 – *"We changed the capacitance in the experiment by separating the foils and by separating the sheets between three pages."*

PostQ2 – *"The greater plate area gives greater capacitance and less plate area gives less capacitance and the further plate spacing gives less capacitance and closer plate spacing gives greater capacitance."*

Thus, network analysis allows us to track changes in, and/assess the sophistication of learner's thinking, in active learning environments, typically employed in introductory university physics courses, and to address in-context, 'issues' or epistemological barriers' which maybe hindering the activation of canonical connections. Students F33 and F14 above, clearly could have benefitted from such interventions.

(c) Mean Epistemic Network summaries by Question

Figures 4.8 and 4.10 present different views (matrix rotations or loadings), of Pre-Post networks. Figure 4.10 below, presents centered networks for each question, unlike Figure 4.8, which was obtained by plotting the combined ENA set of data for PreQ, PostQ1, and PostQ2. The graphs in the figure above, are obtained by creating a separate ENA set for each dataset (PreQ, PostQ1, PostQ2). The variances along the axes are therefore different, whereas in Figure 4.8, they remained the same, and so did the positions of nodes (KEs) in the network projection space. The results are however, not affected; the views are different.

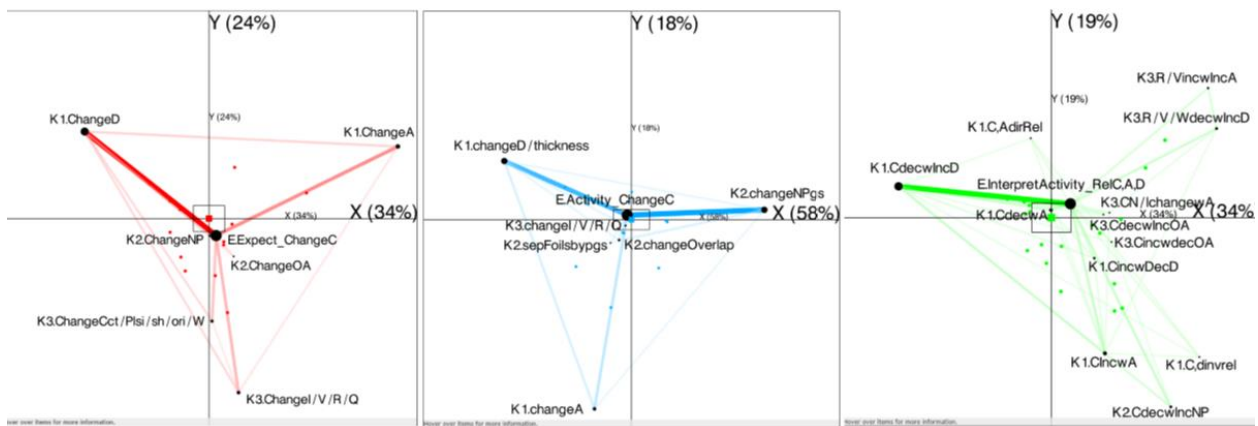


Figure 4.10. Mean epistemic networks of students for PreQ (left graph), PostQ1 (center graph), and PostQ2 (right graph).

As can be seen from Figure 4.10, the node-centrality shifted from canonical before the experiment, to canonical and concrete-intuitive, for PostQ1, and to canonical for PostQ2, after the experiment. Before the experiment, strongest connections were made to the KEs ‘change distance’, and less strong to ‘change area’, which did not change after the experiment, since the students predominantly framed the relationships between physical quantities associated with capacitance, in terms of distance, and demonstrated intuitive rather than canonical knowledge of how it was affected by change in area.

(d) Pre-post Network Model Comparison by Section

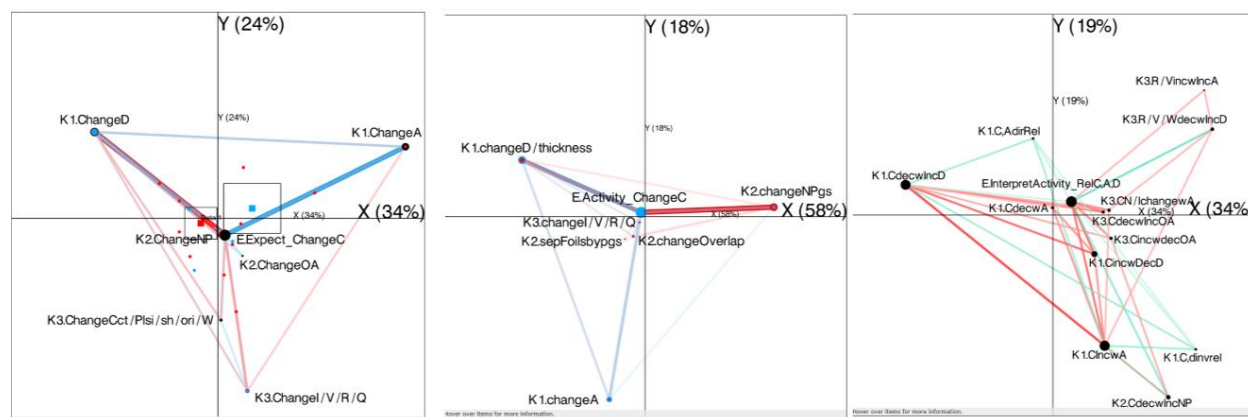


Figure 4.11. Subtracted mean networks of Section 1 (red color) and Section 3 (blue color), for PreQ (left graph), PostQ1 (center graph), and PostQ2 (right graph).

Each network graph in Figure 4.11, highlights the differences in mean networks of students from Section 1 and Section 3, for one question. The graphs were obtained by using the ‘Subtract Equiloards’ feature of ENA webtool. Before the experiment, students of Section 3 were found to frame change in capacitance, in terms of change in both distance and area, while those of Section 1, framed strongly in terms of change in distance. After the experiment, most students of Section 3 did not refer to ‘having changed area’ during the experiment, although they did connect to area before the experiment. Section 1 students, on an average, associated ‘change in distance’ with ‘change in number of pages’, as expected, and gave canonically appropriate descriptions of relationships between capacitance, area, and distance, after the experiment. However, for Post Q2, more connections were made to K3 KEs by Section 1 students, than students of Section 3. It should be noted that, the above graphs show ‘differences’ – those connections of each Section, which were found to be ‘more different’ than the other one, are being presented by the graphs.

At an alpha-level of 0.05, the mean networks of students from Section 1 and Section 3, were found to be significantly different along Dimension 1, for both PreQ and PostQ1 (Table 4.14). As per the norms of statistical significance, the p-values were less than 0.05, values of Cohen’s d were greater than 0.8, suggesting large effect size, and T-statistic values were less than -2 and greater than 2 respectively.

Table 4.14 T-test for Significance of Differences in Epistemic Networks between of Section 1 and Section 3 for PreQ, Post Q1, and PostQ2

	PreQ Dim 1	Dim 2	Post Q1 Dim1	Dim 2	PostQ2 Dim1	Dim2
Mean Section 1	-0.153	-0.044	0.182	0.065	-0.008	0.01
Mean Section 3	0.305	0.089	-0.315	-0.112	0.019	0.023
T statistic	-3.318	-1.065	2.916	1.47	-0.165	-0.271
P value	0.003	0.296	0.006	0.157	0.871	0.789
Cohen's d	-1.138	-0.352	.923	0.548	-0.064	-0.104

(e) Pre-post Network Comparison by Level of Scores on Quantitative Instruments

EBAPS. Results of t-test found no significant differences for PreQ, PostQ1, and PostQ2, between mean epistemic networks of students whose EBAPS scores were low, medium and high (Table 6.1, Appendix 6). Before experiment (Figure 6.1UG, Appendix 6), all 3 groups made strong connections of ‘change capacitance’, to canonical KE, ‘change distance’, but high EBAPS scorers, also connected equally strongly (almost same edge thickness for both KEs) to the canonical KE ‘change area’, unlike medium and high scorers, who made weak connections. PostQ1 network structures (Figure 6.1CG, Appendix 6) were almost identical for the three groups, making strong connections to K1, ‘change d’, and to K2, ‘change number of pages’. However, PostQ2 network structures (Figure 6.1LG, Appendix 6) of medium and high scorers show more connections to abstract-intuitive KEs, than low scorers. Thus, differentiation by EBAPS scores, was found not to be mirrored in epistemic network structures, for this experiment.

A3 HW Quiz. Results of t-test found no significant differences for PreQ, PostQ1, and PostQ2, between mean epistemic networks of students whose A3 HW Quiz scores were low, medium and high (Table 6.2, Appendix 6). Before experiment (Figure 6.2 UG, Appendix 6), mean network structure of high scorers was equivalent to the combined mean networks of low and medium scorers. For PostQ2 (Figure 6.2 UG, Appendix 6), medium and high score groups had more extensive and diverse networks, and these students made many types of K1-K2-K3 connections. For all three questions, the networks were identical for the three groups, on an average, and showed no relation to differences in HW Quiz scores.

Lab4 Quiz. Before the experiment, mean networks of groups scoring high and low on the Lab 4 Quiz, differed significantly, as seen by the results of t-test (Table 6.3, Appendix 6). Comparison of their network graphs (Figure 6.3UG), pointed to differences in the type of K2 connections made – high score group framed the change in capacitance, in terms of ‘change in number of pages’, while the high scores group framed it in terms of ‘change in overlap area’. Low score group networks also showed weaker connections to ‘change in area’, in comparison to high score group. After the experiment, no differences

were found between groups, and their network structures were similar, except for the fact, that medium score group made less K3 connections, unlike low and high score groups, for PostQ2.

Lab Scores. Comparison of network graphs for PreQ (Figure 6.4 UG, Appendix 6) found that all three groups made strong connections to canonical KE, ‘change distance’, less strong to K1, ‘change area’, and weak connections to abstract-intuitive KEs, before experiment. After experiment, the frames of all three groups were still found lacking in strong connections to ‘change area’, although connections to abstract-intuitive KEs were reduced, and some new concrete-intuitive KEs, e.g. ‘change number of pages’ were evidenced in their networks (Figure 6.4 CG, Appendix 6). For deducing the relationships between C, A, and d, (PostQ2), medium score group had many more abstract-intuitive connections, than the other two groups (Figure 6.4 LG, Appendix 6). However, results of t-test found no significant differences for PreQ, PostQ1, and PostQ2, between mean epistemic networks of students whose Lab scores were low, medium and high (Table 6.4, Appendix 6).

Test1 Scores. There were no significant differences found, between the mean networks of Low, Medium, and High scorers of Test 1, as revealed by the results of t-test for PreQ, PostQ1, and PostQ2 (Table 6.5, Appendix 6). Mean networks of all three groups were found to show strong connections to K1, ‘change d’, less strong to K1, ‘change A’ and weak connections to concrete and abstract-intuitive KEs, before the experiment (Figure 6.5 UG, Appendix 6). After the experiment, network structure was found to shift towards reduction of connections to K3 KEs, no change in K1 connections, and introduction of a new concrete-intuitive KE, ‘change no of pgs’ (Figure 6.5 CG, Appendix 6). For PostQ2, while the network structures of PostQ1 were found to have been reproduced, the structures were also found to become more extensive, and included abstract-intuitive ideas about the relationships between C, A and d, with some finer differences in the type of connections made, between groups. Medium and high scorers demonstrated progressively greater sophistication in terms of deducing the relationships both concretely, and mathematically. K3 KEs in the network of ‘high’ group, were also less in comparison to the other two groups, for PostQ2.

Experiment 3 (Ohmic and Non-ohmic Materials)

In this experiment, students explored the graphical relationships between voltage and current in ohmic and non-ohmic materials. PreQ1 and PreQ2, sought to know their prior ideas and expectations about these materials, by asking them to describe ohmic and non-ohmic materials, respectively. Post Q1 and PostQ2 asked them to identify and justify the ohmic or non-ohmic nature of ‘light bulb’ and ‘LED’ respectively, based on their finding from the experiment. This sub-section will present the findings of network analysis, of learners’ coded Twitter responses, to all three questions. As for experiment 2, the network graphs and tables of t-test results, for group comparisons on five quantitative assessment scores,

will be appended in Appendix 7, and discussion will be limited to bringing out similarities and differences between Low, Medium, and High score groups.

(a) Pre-Post Mean Epistemic Network Model Comparison

Table 4.15 T-test for Significance of Differences in Mean Epistemic Networks of students, for PreQ2, PostQ1, and PostQ2

	PostQ1		Post Q2	
	Dim 1	Dim 2	Dim1	Dim 2
Mean PreQ2	-0.052	0.109	-0.052	0.109
Mean Post Q	0.164	0.025	0.149	0.027
T statistic	2.279	-3.051	2.193	-2.921
P value	0.026	0.004	0.031	0.005
Cohen's d	0.512	-0.659	0.492	-0.632

Since the LED and light bulb belonged to the category of non-ohmic materials, the t-test was conducted to test for differences between networks of PreQ2 (non-ohmic) and PostQ1(light bulb), and PreQ2 and PostQ2 (LED). Results (Table 4.15 above) indicate that there were statistically significant structural differences in learners' epistemic networks, which represented their framing of non-ohmic materials, before and after the experiment. At an alpha-level of .05, the p-values were less than 0.05, along both dimensions, and the effect size (the differences in means), as reflected by the values of Cohen's d, was also of medium size (ranging from 0.492 to -0.659).

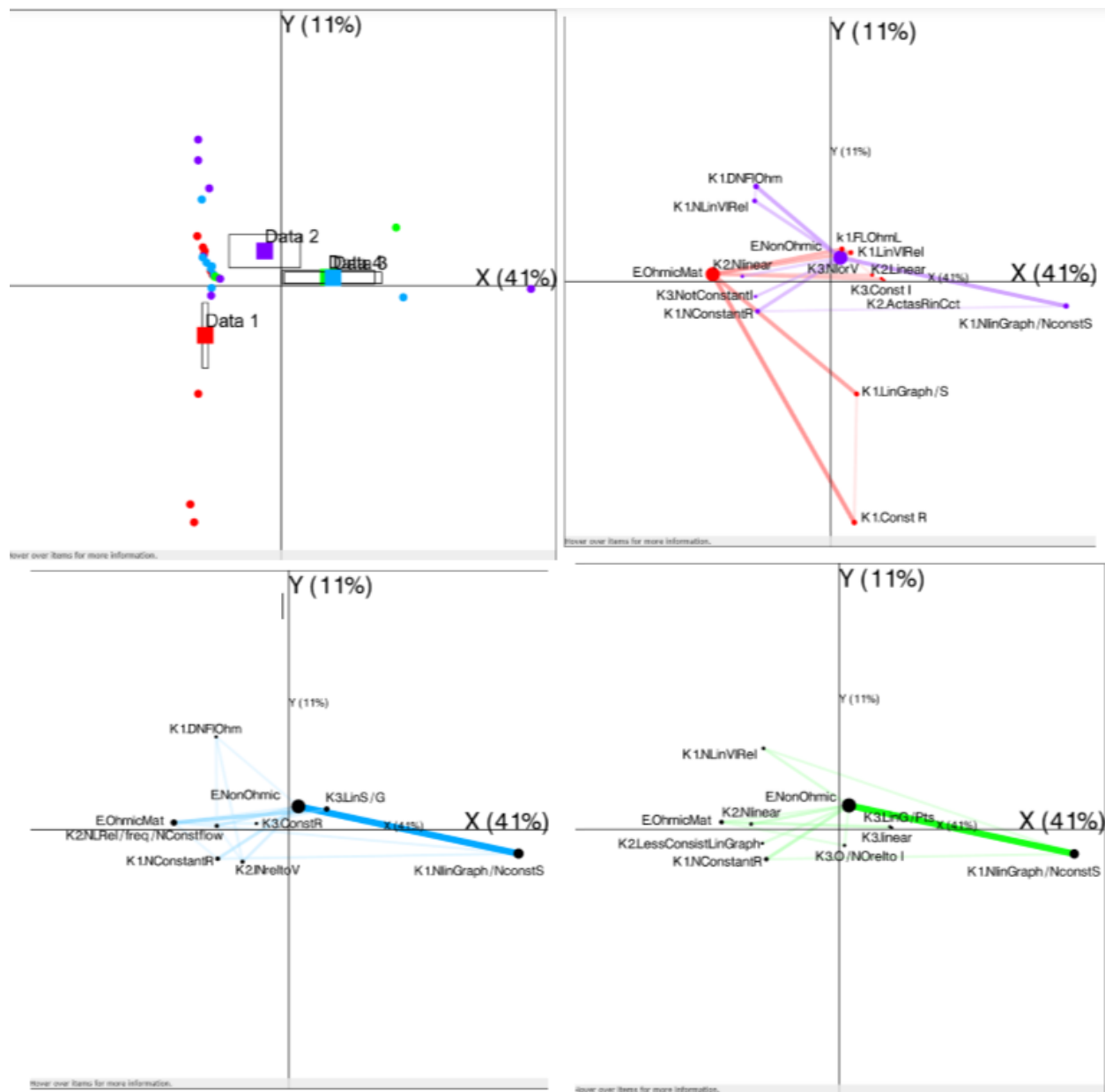


Figure 4.12. Network Centroids of all students for PreQ1, PreQ2, PostQ1, PostQ2 (Upper left graph), and Mean Network graphs of students for PreQ1&2 (Upper right graph), Post Q1(lower left graph), and PostQ2(Lower right graph).

Figure 4.12 presents mean networks (mean of all students' individual networks) of students' framing of PreQ2 (purple color in upper right graph), PostQ1 (lower graph in blue), and PostQ2 (lower graph in green), on the same projection space, meaning that the nodes of knowledge elements occupy the same position in all three graphs, making it easier to track shifts in thinking. Mean (enhanced graphs) networks for each question, will be discussed individually, in sub-section (c) below. The graphs above, reveal that learners framed non-ohmic materials using graphical (non-linear graph/non-constant slope), mathematical (non-linear VI relationship), and verbal (constant R; do not follow ohm's law) representations of canonical knowledge, of nearly equal connection strength, before the experiment. Networks after the experiment, for bulb (blue graph above), and LED (green graph), reflected strong connections to canonical graphical representation, which was to be expected, since the structure of experiment 3, directed learners' attention towards graphical VI relationships. However, the network structures also revealed weak, but newly introduced, concrete and abstract-intuitive knowledge elements, in learners' thinking, e.g. 'less consistent linear graph' (K2), and 'ohmic-non-ohmic related to values of current' (K3), after the experiment. This finding suggested that the experiment, while effective in directing learners' attention to graphs, also confused some of them, possibly because of prior inexperience or misbeliefs about 'graphical interpretation'. Hence, it also supports the argument for ENA-based contextual assessments to track and redirect the misconceptions, or missed connections in learners' thinking.

(b) Pre-Post Individual Student Epistemic Network Examples

Figure 4.13 presents shifts in the epistemic networks of students F8 (ULG), M2(URG), F37(LLG), and F5(LRG) for PreQ1 (red color network), PreQ2 (purple color), PostQ1 (blue), and PostQ2 (green). It was found that the epistemic frame of F8, shifted from canonical framing of ohmic and non-ohmic materials as those which did or did not follow Ohm's Law and did or did not demonstrate linear VI relationship, before the experiment, to concrete intuitive connections to K2.nonlinear for bulb, and to abstract-intuitive connections, K3.ConstantR, and K3.LinearGraph/Slope, for both bulb, and LED. Before the experiment Hence clearly, it was the graphs which caused the students' conceptual clarity before the experiment, to be diffused by her difficulty in reconciling her prior notions of 'non-linearity' of graph, with what she found from the experiment. Her specific responses were as follows:

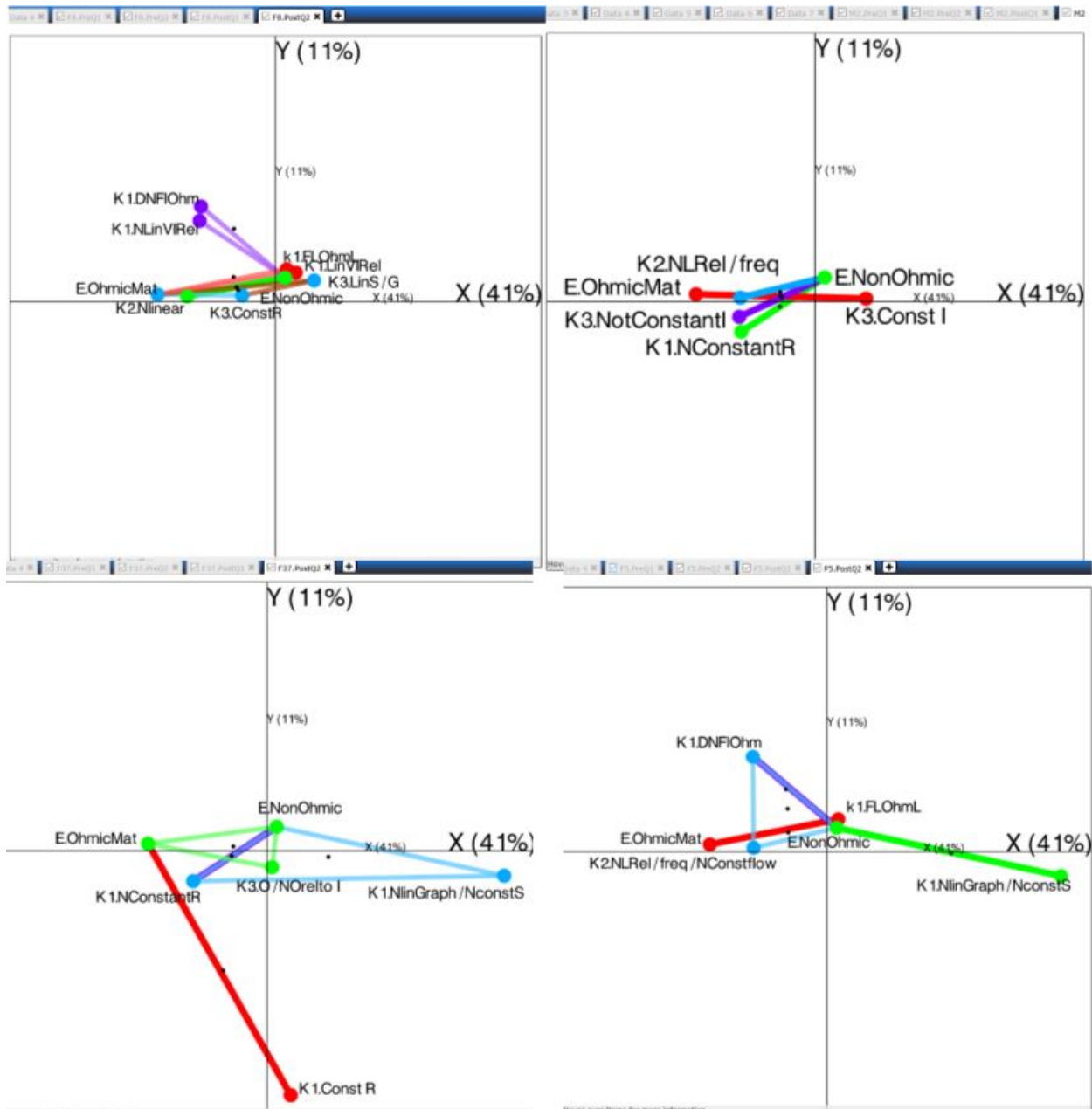


Figure 4.13. Network models of student F8 (upper left graph), M2(upper right graph), F37(lower left graph), and F5(lower right graph) for PreQ1(red), PreQ2(dark blue), PostQ1(light blue), and PostQ2 (green).

PreQ1and2 – *“the difference is whether or not they follow ohm's law. An ohmic conductor would have linear relationship between the current and voltage”*

PostQ1- *“The bulb is a ohmic resistor because it does have a constant resistance... a linear slope”*.

PostQ2 – *“The LED is not an ohmic resistor because it is not linear”*

The epistemic frame of M2 on the other hand, was found to shift from abstract-intuitive KEs, K3.NonConstant I (non-ohmic materials), and K3.constant I (ohmic materials), before the experiment, to concrete-intuitive KE, K2.NonLinearRelation/Frequency, for bulb, and canonical KE, K1.Non Constant R, for LED. Thus, his prior abstract notions of ohmic and non-ohmic materials, seemed to have been redirected through the concrete experience of his observations, towards framing them in terms of graphical VI relationship. His specific responses were as follows:

PreQ – *“ohmic materials create a constant linear current and non ohmic materials do not”*

PostQ1 – *“the bulb is not a ohmic resistor because it does not create a constant linear frequency”*

PostQ2 – *“the LED is not a ohmic resistor either because it isn't constant but did seem to have a more consistent frequency than the bulb”*

The frame of F37, like F8, was found to shift from canonical knowledge of ohmic and/ non-ohmic materials as those which have or do not have constant resistance, or which generate a non-linear graphical output and/ non-constant slope, before the experiment, to abstract-intuitive interpretation, that LED's ohmic or non-ohmic behavior was related to the value of current. Her specific responses were as follows:

PreQ – *“Ohmic materials have constant resistance and non-ohmic materials do not”*

PostQ1 – *“The bulb is not an ohmic resistor because its graph of voltage versus current does not have a constant slope which represents resistance”*.

PostQ2 – *“The LED is an ohmic resistor only at certain currents. At the current 0.004A the LED has an undefined slope and therefore is not an ohmic resistor at that current”*.

In this case, the student did seem to have appropriate canonical ideas, but the nature of graphical output of LED, was clearly, not what she could reconcile as being ‘non-linear’, either because she had prior ‘notions’ which hindered her reasoning, or simply lack of experience with interpreting non-linear graphs, because her interpretation of the graph was quite specific, to the point, and based directly on her observations. Student F 5's frame was found to shift from canonical KE before experiment, to concrete-intuitive KE for bulb, after experiment, for similar reasons, except she abstracted correctly based on the concrete findings of her observations. Her responses are as follows:

PreQ – *“ohmic materials follow ohm's law while non- ohmic materials do”*

PostQ1 – *“the bulb is a non- ohmic resistor because it does not follow ohms and its does not represent linear data”*

PostQ2 – *“the LED is also a non- ohmic resistor because of not following ohms law and the slope is not linear”*

These results demonstrate that network analyses present a viable option, for tracking shifts in individual learner's thinking, and therefore, for assessing the sophistication of their thinking.

(c) PreQ and PostQ Mean Network Models

Network graphs of Figure 4.14 provide an aggregated representation of learners' thinking for each of the four questions they were asked to respond to. Students made most connections to graphical, mathematical, and verbal representations of canonical knowledge to describe ohmic and non-ohmic materials before the experiment (Upper graphs, Figure 4.14). Abstract-intuitive KEs were related to descriptions of current remaining constant for ohmic, and non-constant, for non-ohmic, instead of resistance. Abstract-intuitive KEs were related to 'incomplete, somewhat intuitive' description of ohmic materials as 'linear', and non-ohmic as 'non-linear'.

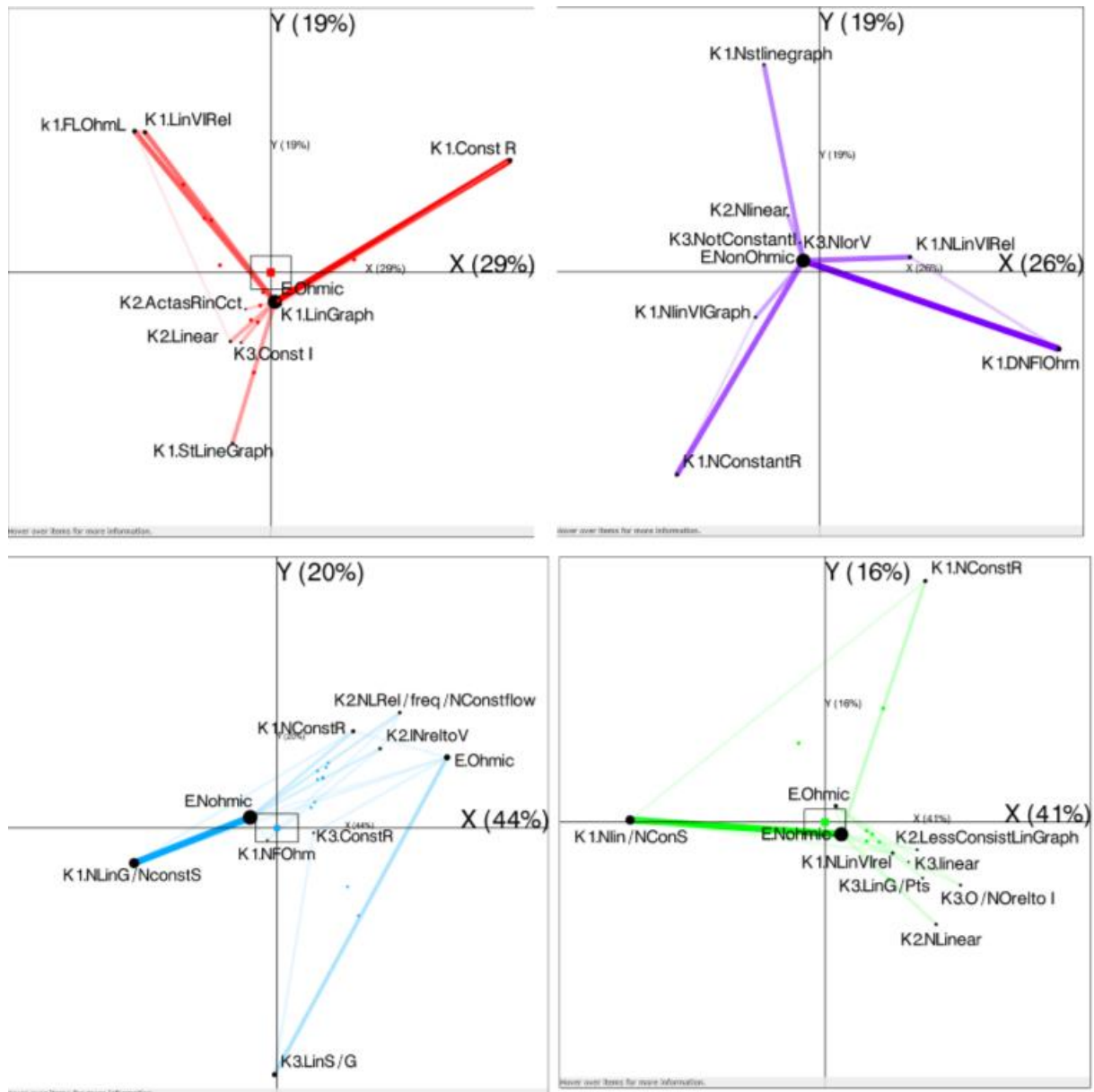


Figure 4.14. Mean Network models of all students for PreQ1 (upper left graph), PreQ2 (upper right graph), PostQ1(lower left graph), and PostQ2 (lower right graph).

Most students identified bulb as non-ohmic, because of non-linearity of graph, and/non-constancy of slope of graph. Thus, graphical representations of canonical knowledge dominated learners' framing of bulb, as well as LED, as non-ohmic materials. Students who framed LED and/ bulb as ohmic materials, were also found to frame the construct in terms of graph, and/ slope, but their interpretations seemed to be influenced by prior misbeliefs or weak formulations about non-linearity and linearity of graphs. (Figure 4.14, Lower graphs). These inferences corroborate, and therefore validate, the findings derived from other Pre-post network graph analyses of Experiment 3, presented above and in the following sub-sections, as well as, qualitative findings of Experiment 3 in the previous section.

(d) Pre-post Network Comparison by Section

Table 4.16 T-test for Significance of Differences in Epistemic Networks of students by Section of Course [PreQ2, Post Q1, and PostQ2]

	PreQ2 Dim 1	Dim 2	Post Q1 Dim1	Dim 2	PostQ2 Dim1	Dim2
Mean Section 1	-0.182	0.146	0.006	0.035	0.121	0.036
Mean Section 2	0.192	0.04	0.427	0.007	0.196	0.01
T statistic	-2.791	2.341	-2.817	1.766	-0.483	1.52
P value	0.013	0.024	0.009	0.085	0.633	0.14
Cohen's d	-1.184	0.659	-0.979	0.52	-0.167	0.402

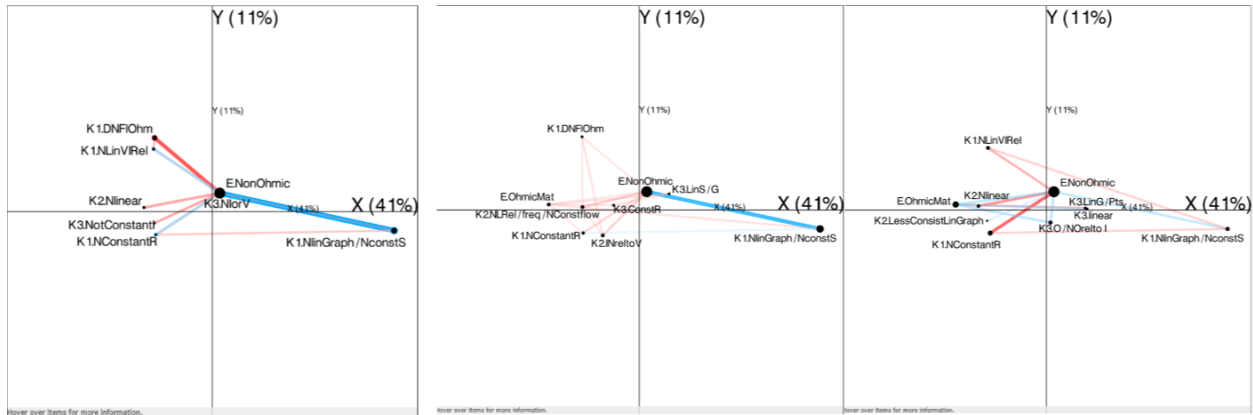


Figure 4.15. Subtracted Mean Network models of students in Section 1 and Section 3, for PreQ2 (left graph), PostQ1(center graph), and PostQ2(right graph).

Results of t-test, at an alpha-level of .05, revealed that there were significant differences between the network connections of students from Section 1 and Section 3 of the course, along both dimensions, with respect to how they framed non-ohmic materials before the experiment, and along Dimension 1, with respect to how they framed bulb as a non-ohmic material, after the experiment. (Table 4.16 above). The p-values were below .05, and therefore, the ‘default’ null hypothesis of t-test, that ‘there were no differences between means of the two groups’ was rejected. The effects sizes (Cohen’s d) were medium (greater than 0.5) to strong (greater than 0.8).

The graphs of Figure 4.15, were derived by subtracting mean networks of sections 1(red) and 3 (blue), for PreQ2, PostQ1, and PostQ2. The connections remaining, highlight the structural differences between the networks of both sections. Before the experiment, more students of Section 3 framed non-ohmic materials using graphical, mathematical and verbal canonical knowledge representations, while more students of Section 1 made intuitive connections (K2 and K3), as compared to the other section (left graph, Figure 4.15). For the bulb (PostQ1), more students of Section 3, framed bulb as non-ohmic, based on non-linearity of graph, and/non-constancy of slope, while more students of section 1 framed the concept using canonical KEs, ‘constant R’ and ‘do not follow ohm’s law’, along with intuitive KEs (center graph, Figure 4.15).

While the differences in framing the concept of LED as non-ohmic(PostQ2), were not significant, more students of Section 3, framed the concept in terms of abstract -intuitive KE, K3. Ohmic/Non-ohmic related to current’, and to concrete-intuitive KE, K2.less consistent linear graph, than students of Section 1. At the same time, more students of Section 1, framed it in terms of canonical KE, K1.ConstantR, K1.LinearVI relation, and concrete-intuitive KE, K2. Nonlinear.

These findings demonstrate the viability of using ENA for comparing shifts in thinking, by Sections of a course, which could further serve as a useful evaluation mechanism of instruction, for the TAs, and the instructor.

(e) Pre-post Network Comparison by Level of Scores on Quantitative Instruments

EBAPS. Although no significant differences were found between the epistemic networks of students grouped by low, medium and high scores on the EBAPS survey (Table 7.1, Appendix 7), examination of network graphs of the three groups for PreQ2, revealed that while all three groups of students connected predominantly to canonical KEs, of different types, medium scorers made weak connections to abstract-intuitive, and high scorers, to concrete-intuitive KEs before the experiment (Figure 7.1UG, Appendix 7). However, after the experiment, epistemic frames of all three groups appeared identical for light bulb, with strong connections to graphical justifications of bulb as non-ohmic, and weaker connections to intuitive knowledge elements (K2 and K3) (Figure 7.1CG, Appendix 7). The

results for PostQ2 (LED), were like light bulb, with respect to canonical connections, but the low scorers made no intuitive connections, compared to medium and high scorers, who did (Figure 7.1LG, Appendix 7).

A3 HW Quiz(Attempt1). Significant differences were found between epistemic networks of low and medium score groups, and, medium and high score groups, before the experiment, along Dimension 1 (Table 7.2, Appendix 7). Significant differences were also found, between networks of low score and medium score groups, and, high score and low score groups, for PostQ1(light bulb), along Dimension 2. At an alpha-level of .05, p-values for these tests were less than .05, and corresponding effect size was large (Cohen's d between 0.7 and 2.975). Examination of network graphs for PreQ2 revealed that different types of connections were made by medium scorers, as compared to low and high scorers (Figure 7.2UG, Appendix 7). Medium scorers made strong connections to canonical knowledge that non-ohmic materials did not follow ohm's law, and demonstrated non-linear VI relations. In contrast, low and high scorers connected strongly to non-linear graphical relation/graph, and non-constant resistance.

For PostQ1 (Figure 7.2CG, Appendix 7), structural differences in networks were evident, with low scorers having no connections to abstract-intuitive KEs, and making more diverse canonical connections, compared to both medium, and high scoring groups. High scoring group, on the other hand, were found to make diverse intuitive connections (K2 and K3). While the same trend continued for Post Q2 (LED), the structural differences were less striking, as corroborated by the t-test results which showed no significant differences for the LED question (Figure 7.2LG, Appendix 7).

Lab 4 Quiz. Results of t-test indicated significant differences between epistemic networks of low and high score groups, along Dimension 2 for PreQ2, and PostQ1, and along Dimension 1, for PostQ2. (Table 7.3, Appendix 7). Significant differences were also found, between networks of medium score and high score groups, along Dimension 1, for both PostQ1, and PostQ2. Significant differences between low and medium score groups were found along Dimension 2, for PostQ1, alone. At an alpha-level of .05, p-values for these tests were less than .05, and corresponding effect size was large (Cohen's d between 0.859 and 1.503).

For non-ohmic materials (PreQ2), low scorers thought in terms of 'resistance', and 'ohm's law', medium scorers thought in terms of 'graph', in addition to 'resistance', and 'ohm's law', and high scorers's thinking excluded reference to 'resistance', but included strong connections to 'graph', 'ohm's law', and 'VI relationship'. For PostQ1(light bulb), high score group made strong canonical connections to K1.NLinGraph/NConstSlope, medium score group made diverse connections to K1,K2,and K3, while low score group's connections were diverse, but excluded K3. Post Q2 networks followed the same

pattern as PostQ1, but low score group's connections included abstract-intuitive, and excluded concrete-intuitive connections, unlike PostQ1.

Lab 4. Results of t-test indicated significant differences between epistemic networks of low and high score groups, along Dimension 1 for PreQ2, and PostQ1. At an alpha-level of .05, p-values for these tests were less than .05, and corresponding effect size was large (Cohen's d between 0.864 and 1.077). Before the experiment, low score group made strong connections to canonical knowledge that non-ohmic materials did not follow ohm's law; high score group made more connections to 'constant resistance' and 'nonlinear graph', while medium score group made diverse connections of nearly equal strength, to K1, K2, and K3 (Figure 7.4UG, Appendix 7).

For the 'light bulb' question, low score group made most connections to concrete-intuitive KE, 'nonlinear relation/frequency', and weaker connections to other K1, K2, and K3 knowledge elements. Medium and high score groups connected strongly to canonical KE, 'nonlinear graph/non-constant slope', and made lesser (than low group), weak connections to K2, and K3. Also, more medium scorers were found to have framed light bulb, as ohmic, as compared to low and high score groups (Figure 7.4CG, Appendix 7). For PostQ2 (LED), all 3 groups, connected strongly to K1, nonlinear graph/nonconstant slope, and less students framed the LED as ohmic, as compared to light bulb (Figure 7.4LG, Appendix 7).

Test1. Results of t-test indicated significant differences between epistemic networks of low and medium score groups, along Dimension 1 for PreQ2 (Table 7.5, Appendix 7). Significant differences were also found, between networks of medium score and high score groups, along Dimension 1, for PostQ2. At an alpha-level of .05, p-values for these tests were less than .05, and corresponding effect size was large (Cohen's d between 0.891 and 1.012). Before the experiment, low score group were found to make strong connections to canonical knowledge that non-ohmic materials did not follow ohm's law, medium score group to 'nonconstant resistance', and 'nonlinear graph', and high score group revealed no dominant connection. All 3 groups revealed weak connections to other K1, K2, and K3 (Figure 7.5UG, Appendix 7).

For the 'light bulb' question, low score group made strong connections to concrete-intuitive KE, 'nonlinear relation/frequency', and all 3 groups connected strongly to canonical KE, 'nonlinear graph/non-constant slope'. Weaker connections were revealed to other K1, K2, and K3, by the three groups. Also, more high scorers were found to have framed light bulb, as ohmic, as compared to low and high score groups (Figure 7.5CG, Appendix 7). For PostQ2 (LED), all 3 groups, connected strongly to K1, nonlinear graph/nonconstant slope, and less students framed the LED as ohmic, as compared to light bulb (Figure 7.5LG, Appendix 7).

Experiment 4 (RC Circuits)

In this experiment, students investigated the behavior of voltage and current in RC circuits, through graphical outputs generated using Pasco RLC interface. PreQ sought to know their prior ideas and expectations about how the behavior of current would be affected when a capacitor was introduced in a circuit containing only resistance. Post Q1 asked them to explain what they had learned about the behavior of current in RC circuits, from the experiment, and PostQ2 asked them to compare the behaviors of voltage and current, as per their findings from the experiment. This sub-section will present the findings of network analysis, of learners' coded Twitter responses, to all three questions. The network graphs and tables of t-test results, for group comparisons on five quantitative assessment scores, will be appended in Appendix 8, and discussion will be limited to bringing out similarities and differences between Low, Medium, and High score groups.

(a) PreQ and Post Q Mean and Individual Student Epistemic Networks

T-test for examining differences between PreQ and Post Q networks, was conducted only for PostQ1, since PostQ2 was conceptually different from PreQ, and therefore, differences were bound to be found, between the epistemic networks of PostQ2 and PreQ, given the fact that the networks were derived from independent sets of codes (knowledge elements). The results indicated that there were significant differences between mean epistemic networks pertaining to student responses to questions about 'the behavior of current in RC circuits', before and after the experiment (Table 4.17 below). At an alpha-level of .05, the p-value was found to be less than .05, along both dimensions, and effect sizes were medium (Cohen's $d = 0.571$), to large (Cohen's $d = 0.881$), thus confirming statistically, that the alternative hypothesis that there were differences, be accepted.

Table 4.17 T-test for Significance of Differences in Mean Epistemic Networks of students for PreQ, and Post Q1

	Dim 1	Dim 2
Mean PreQ	0.178	-0.065
Mean PostQ1	-0.087	0.086
T statistic	3.669	-2.264
P value	<.001	0.03
Cohen's d	0.831	-0.571

Figure 4.16 below, presents the graphs of mean epistemic networks centroid positions of epistemic networks, and individual students' epistemic networks for all three questions, positioned as explained in the description of figure 4.16.

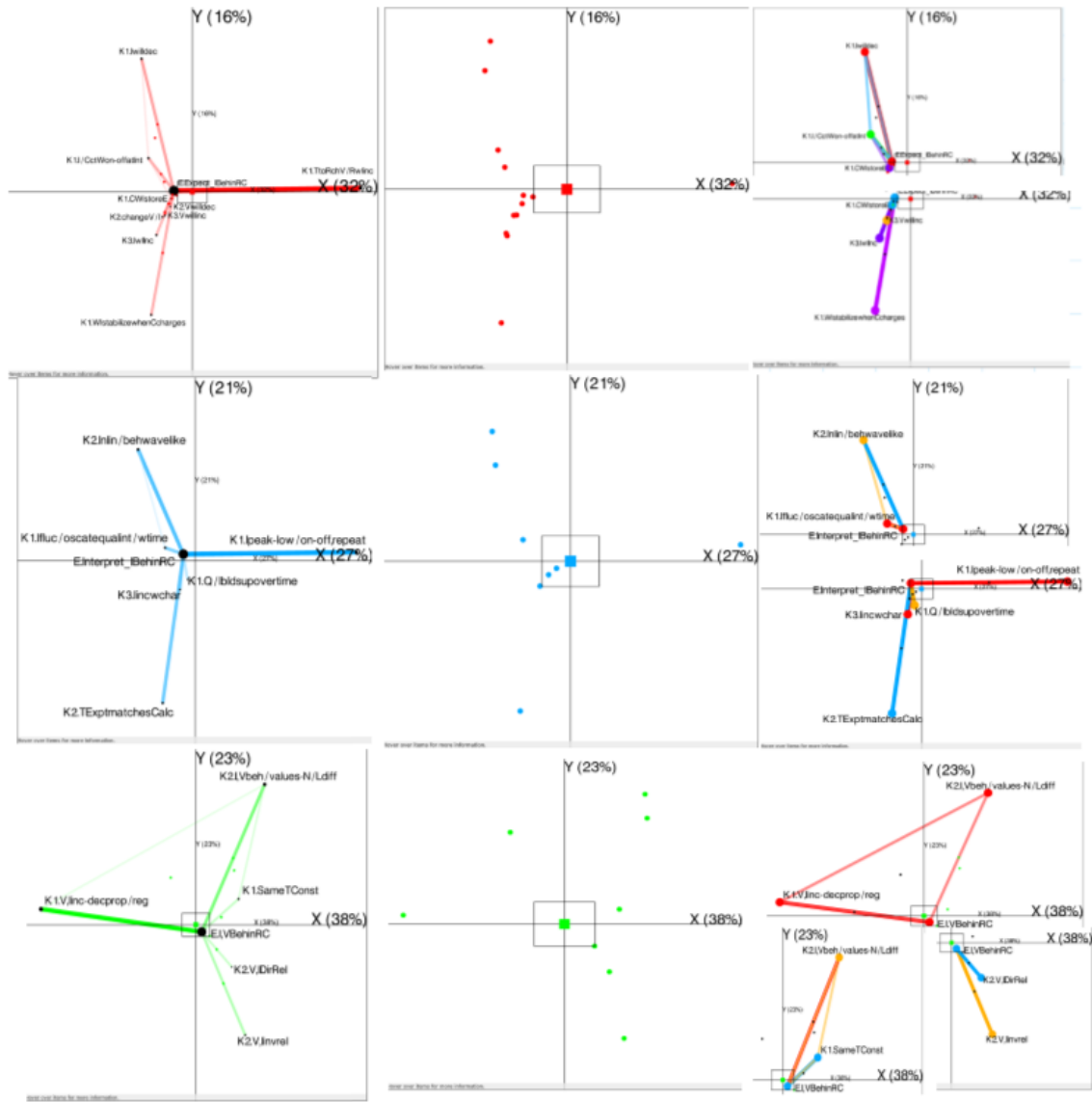


Figure 4.16. Mean Networks of students (left graphs), and Centroid positions of epistemic networks of all students (center graphs) for PreQ (red), PostQ1(blue), and PostQ2(green); Individual student models (right graphs, different color for each student) for PreQ.

Before the experiment, students framed the behavior of current, predominantly using canonical explanations, ‘time to reach maximum value will increase’ (Figure 4.16ULG). They also used other canonical KEs, ‘current will decrease’, ‘current will stabilize when capacitor charges’, and intuitive KEs referring to ‘change in voltage, current, or charge’, and ‘increase in voltage or current’, to a lesser extent, as revealed by the graph. After the experiment their frames shifted towards graphical canonical and concrete-intuitive explanations, of ‘current reaching peaks and lows, repeatedly’(K1) and ‘behaving wavelike’ (K2), or ‘time constant of current obtained experimentally matching calculated values’ (K2) (Figure 4.16 CLG). Some student’s thinking, however, did retain connections to K3, ‘current increase with charging’. For PostQ2, most connections were found to have been made to K1, ‘voltage and current increase and decrease regularly, or proportionally’, and to K2, that behaviors of voltage and current were similar, or that they revealed either direct, or inverse relationships (Figure 4.16 LLG). The latter KE was concrete, based on the source – graph, and intuitive, based on incomplete interpretation of the graphical output.

For PreQ, students whose network centroids were found to be in quadrant 2 of graph (Figure4.16 UCG), were found to make connections to K1, ‘current will decrease’, and ‘current/circuit will on-off at intervals (Figure 4.16 URG). It can also be seen that these two nodes were in quadrant 1 of left upper graph of Figure 4.30, which represents the mean epistemic network for PreQ. Thus, by cross referencing the graphs of centroid positions and mean networks, epistemic networks of individual students (right graphs of Figure 4.16) can be predicted approximately. The graphs above were presented to illustrate how such analyses could be conducted. Thus, the top 3 graphs, center 3 graphs, and bottom 3 graphs have a one-to-one correspondence with one another, in terms of node positions on network graphs (left and right graphs) and centroid positions on the center graphs. Students whose centroids were found to be in quadrant 3 of upper center graph, were found to make connections to K1, ‘current will stabilize when capacitor charges’, and to K3, ‘current and/voltage will increase’ (upper right graph, quadrant 3).

Examining the center 3 graphs for PostQ1, it was found that students whose centroids were positioned in quadrant 2 of center graph, made most connections to K1, ‘current fluctuates/oscillates at regular intervals of time’, and to K2, ‘current was nonlinear/wavelike’ (center right graph, quadrant 2). Whereas students with centroids in quadrant 3 of center (blue) graph, were found to have made most connections to K1, ‘current reaches peaks and lows repeatedly’, or ‘builds up over time’ and to K2, ‘time constants for current match’ (center right graph, quadrants 1 and 3). For Post Q2, centroids were found to be distributed in quadrants 1, 2, and 4, of lower center graph (green), and corresponding student networks (each color represents one student’s network) were found to be in quadrants 1, 2 and 4 of lower right graph (split into quadrants). Quadrant 2 centroids were found to connect most with K1, ‘current and

voltage increase and decrease regularly’, and to K2, ‘current and voltage behave similarly’; quadrant 1 centroids made connections to K1, ‘current and voltage have same time constant’, and to K2, ‘current and voltage behaviors are similar’; and quadrant 4 centroids connected to K2, ‘voltage and current have direct relation’, or ‘inverse relation’.

(b) Pre-Post Individual Student Epistemic Network Examples

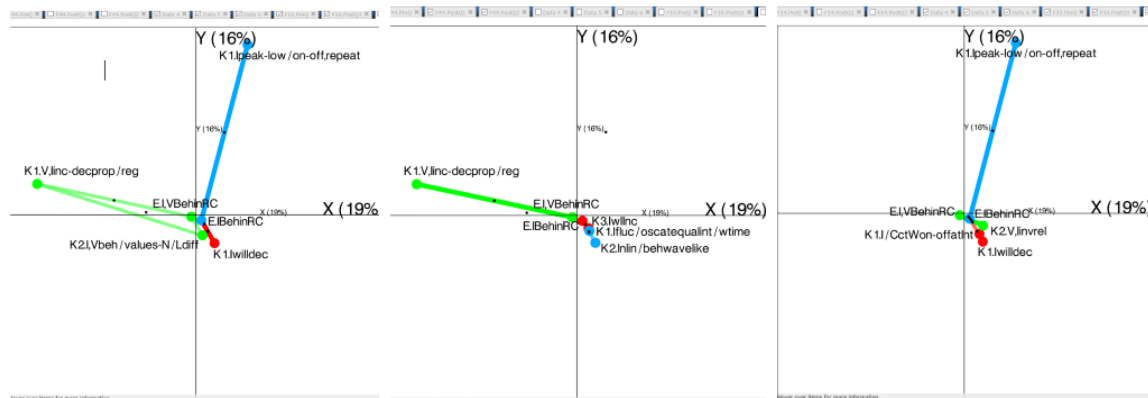


Figure 4.17. Individual networks of students F10 (left graph), F44 (center graph), and F 33(right graph), for PreQ(red), PostQ1(blue), and PostQ2(green).

This subsection presents examples of shifts in thinking, of three students, as they responded to PreQ before the experiment, and to PostQ1, and PostQ2, after the experiment. Before and after the experiment, student F10 framed behavior of current in RC circuits, canonically, as ‘current will decrease’, and ‘current reaches peaks and lows repeatedly’ respectively. Thus, after the experiment, she made strong connection to graphical representation of canonical knowledge. Her comparison of behaviors of voltage and current after the experiment, referred to canonical, and concrete-intuitive KEs, ‘current and voltage increase/decrease regularly’, and ‘current and voltage behave similarly’, respectively. Her responses were as follows:

PreQ – “when capacitor is introduced the current decreases”

PostQ1 – “The current reaches a continuous repeated peak and low.

PostQ2 – “they both behaved the same with continuous repeated peaks and lows”

The epistemic network of F 44, showed connection to abstract-intuitive KE, ‘current will increase’, before the experiment; shifted to K1, ‘current fluctuates with time’, and K2, ‘current behaves wave-like’ for PostQ1; and then to K1, ‘voltage and current increase/ decrease proportionally/regularly’, for PostQ2. Her responses were:

PreQ – “*I think the current will be stronger but I am not entirely sure on this one*”.

PostQ1- “*It had a equal frequency between each peak. Also, the peaks went up at a certain time each time.*”

PostQ2 – “*The similarities were the both had a constant time that they peaked. A difference is the graph had more of a consistency in the current as in the graph was more equal compared to the other graph that did not have linear lines*”.

Student F33’s network shifted from K1, ‘current will decrease’, and ‘current will on-off at intervals’ before the experiment, to graphical K1, ‘current reaches peaks and lows repeatedly’, after the experiment, and then to the concrete-intuitive inference that, ‘current and voltage show inverse relation’. Her responses were:

PreQ – “*the current will change because the capacitor will cause the current to decrease. This is because it breaks up the circuit's constant flow*”

PostQ1 – “*The current reaches the same repeated peak and low*”.

PostQ2 – “*They have a negative relationship, when one goes up, the other goes down*”.

Epistemic network comparisons by groups will be presented in the following subsections, using superimposed graphs of groups, by question. Interpretation will be supported by using centroid graphs to cross-reference the group-positions. Thus, the number of graphs will be limited to six, per analysis, instead of nine graphs used in previous analyses. T-test results will be reported only when significant differences were found.

(c) Pre-post Network Comparison by Section

As seen from the lower three graphs of Figure 4.18 below, the confidence intervals (bigger 2 squares in each graph) of means (solid small squares) of Section 1 (red), and section 2 (blue), overlap significantly, for PreQ (lower left graph), PostQ1(lower center graph), and PostQ2 (lower right graph), suggesting little or no significant differences, between mean networks of students from Section 1 and Section 3 of the course. The t-test results confirmed the above findings (Table 4.15).

Table 4.18 T-test for Significance of Differences in Mean Epistemic Networks of students from Section 1 and Section 3, for PreQ, Post Q, and PostQ2

	PreQ		PostQ1		PostQ2	
	Dim 1	Dim 2	Dim1	Dim 2	Dim1	Dim2
Mean Section1	0.154	0.046	-0.082	0.105	-0.111	-0.145
Mean Section 2	0.224	0.028	-0.098	0.144	-0.114	-0.203
T statistic	-0.439	0.567	1.394	-1.11	0.122	0.373
P value	0.665	0.576	0.177	0.28	0.904	0.714
Cohen’s d	-0.159	0.202	0.519	-0.43	0.049	0.15

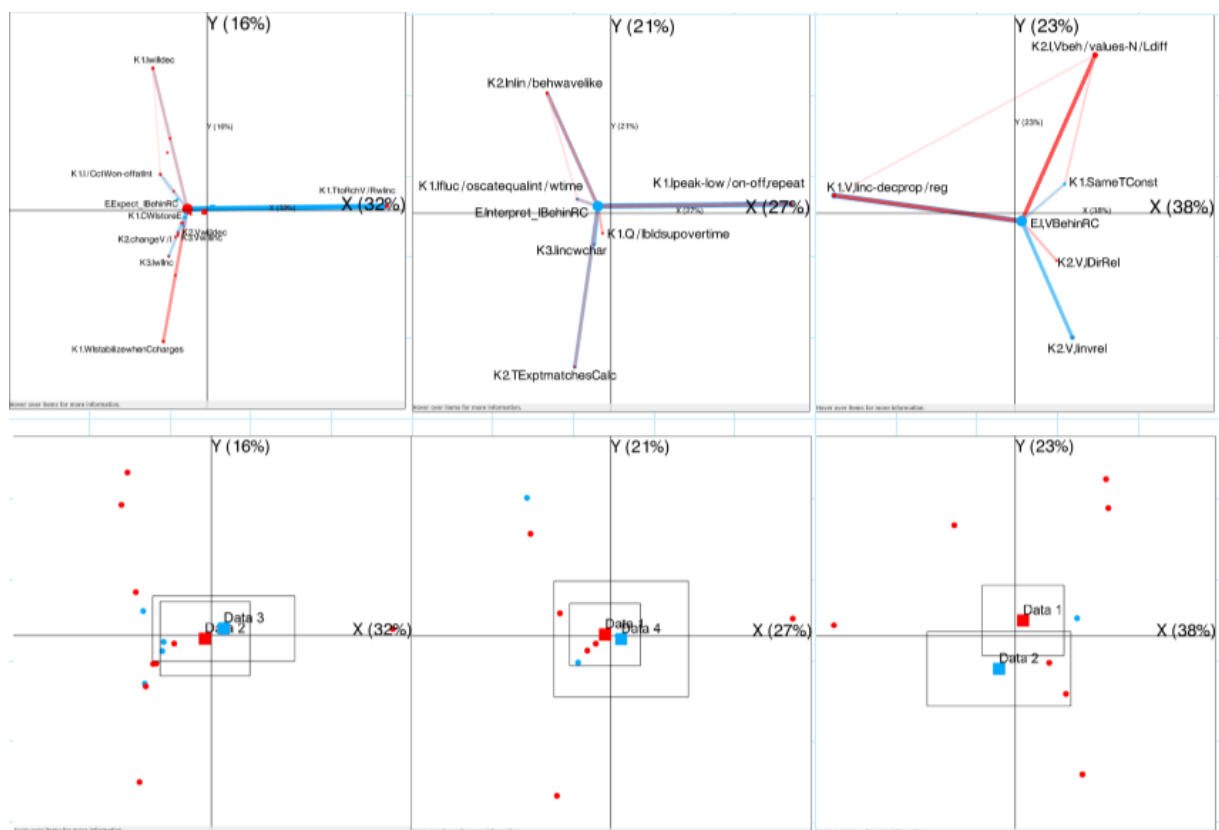


Figure 4.18. Mean epistemic networks of students from Section1(red), and Section 3(blue) (upper 3 graphs), and centroid positions of all students from Section 1(red), and Section 3(blue) (lower 3 graphs), for PreQ (left graphs), PostQ1(center graphs), and PostQ2 (right graphs)

Students of both sections were found to make strong connections to canonical KEs, and weak connections to intuitive KEs before the experiment (upper left graph). Section 1 students made more connections to K1, ‘current will stabilize when capacitor charges’, while Section 3 students made more connections to K1, ‘time to reach maximum value will increase’, than the other section. Students of both sections made strong canonical connections to ‘current will decrease’. After the experiment, students of both sections were found to make strong connections to K1, ‘current repeatedly reaches peaks and lows’, and to K2, ‘experimental and calculated time constants of current match’, or ‘current is nonlinear/wave-like’. For PostQ2, Section 1 students made more connections to K2, ‘current and voltage behavior is similar’, and section 3, to K2, ‘voltage and current show inverse relation’, while students of both Sections made strong connections to K1, ‘voltage and current increase/decrease proportionally, and/regularly’. Differences in K2 connections, framing the similarities and differences in behavior of current and voltage, could have originated from TA differences in explaining the graphs of voltage and current, to students, who asked for their assistance. Perhaps Section 3 TAs, drew students’ attention towards ‘proportional changes’ while Section 1 TAs drew their attention towards the inverse waveforms.

(d) Pre-post Network Comparison by Level of Scores on Quantitative Instruments

EBAPS. Results of t-test indicated that the mean networks of Medium and High score groups differed significantly, along both dimensions for PreQ; and Low and Medium score groups differed significantly along Dimension 2, for PostQ1 (Table 8.1, Appendix 8). The confidence intervals as seen on the network graphs (Figure 8.1LG, Appendix 8), reflect these findings. Before the experiment, although all three groups framed the behavior of current in RC circuits, using canonical KE, ‘current will decrease’, High score group made more connection to K1, ‘time to reach maximum value will increase’, Medium group, to ‘current/circuit will on-off at intervals’, and Low and high groups to ‘current will stabilize when capacitor charges’. The latter abstraction of Low and High score groups could be construed as more sophisticated, since the behavior of current is explicitly connected to what happens with capacitor. After the experiment, Low group was found to make stronger concrete-intuitive connections, while Medium and High groups connected to canonical KE, ‘current reaches peaks and lows repeatedly’, and to K2, ‘observed and calculated values of time constants were same’ (Figure 8.1 UCG, Appendix 8). In response to PostQ2, High group made more diverse connections to K1 and K2. Both Low and High groups connected strongly to K1, ‘current, and voltage increase/decrease proportionally’, while Medium group made more connections to concrete-intuitive KE, ‘behaviors of current and voltage were similar’. High scorers, further observed that ‘current and voltage had inverse relationship’ (Upper right graph of Figure 8.1).

A3HW Quiz. Results of t-test found significant differences between Low and Medium score groups, along both dimensions for PreQ, and along Dimension 2 for PostQ2 (Table 8.2, Appendix 8). Before the experiment, Low group predominantly framed the behavior of current in RC circuits in terms of K1, ‘current will stabilize when capacitor charges’, while Medium and High score groups framed it predominantly in terms of K1, ‘time to reach max value will increase’ (Figure 8.2 ULG, Appendix 8). All 3 groups thought less strongly in terms of K1, ‘current will decrease’, and made weaker connections to other K1, K2, and K3 knowledge elements. For Post Q1, all three groups’ thinking shifted towards K1, ‘current reaches peaks and lows repeatedly’, and K2s, ‘observed and calculated values of time constants of current match’ and ‘current is nonlinear/wavelike’ (Figure 8.2 UCG, Appendix 8). However, Low group also made comparatively more connections to abstract-intuitive KE, ‘current increased with charging’. For PostQ2, Medium and High groups connected strongly to K1, ‘current and voltage increase-decrease proportionally’, while Low and High groups connected strongly to concrete -intuitive KEs, ‘current and voltage have direct and/inverse relation’ (Figure 8.2 URG, Appendix 8). Medium and High groups were also found to connect strongly to K2, ‘behaviors of current and voltage are similar’, while Low group was found to make weaker connections to this KE.

Lab Quiz. At an alpha-level of .05, significant differences were found between mean epistemic networks of Low and High score groups of Lab Quiz, along both dimensions, for PreQ (Table 8.3, Appendix 8). The p-values were found to be less than .05, and Cohen's d values indicated large effect size (greater than 0.8). For PreQ, Low and Medium score groups were found to make strong connections to canonical KE, 'current will stabilize when capacitor charges', while High score group made most connections to K1, 'time to reach maximum value will increase'. All 3 groups also connected to K1, 'current will decrease', and made weaker connections to other abstract-intuitive KEs (Figure 8.3 ULG, Appendix 8). After the experiment, Low and High groups made more connections to concrete-intuitive than to canonical KEs. Low group made most connections to K2, 'current is nonlinear/wavelike', and next to K1, 'current fluctuated at equal intervals', while High group made most connections to K2, 'observed and calculated time constants of current match', and next, to K1, 'current reaches peaks and lows repeatedly', which was also found in the network of Medium group, along with K2, 'current is nonlinear' (Figure 8.3 UCG, Appendix 8). For PostQ2, Low group was found to make most connections to K1, 'voltage, current increase-decrease proportionally', and to K2, 'V, I behavior similar', Medium group, to concrete-intuitive KEs, that voltage and current have 'direct', and/ 'inverse' relation, while High group shared common K2 connection, 'inverse relation' with Medium group, and K1 connection, 'I peaks-lows' with Low group (Figure 8.3 URG, Appendix 8).

Lab Scores. At an alpha-level of .05, significant differences were found between mean epistemic networks of Low and High score groups of Lab score, along both dimensions, for PostQ1 (Table 8.4, Appendix 8). p-values were found to be less than .05, and effect sizes were large (Cohen's d values greater than 0.8). For PreQ, Low and High score groups were found to make strong connections to canonical KE, 'time to reach maximum value will increase', while Medium and High score groups shared common connections to K1s, 'current will decrease', and 'I will stabilize when C charges' (Figure 8.4 ULG, Appendix 8). After the experiment, Low and High groups shared common connections to concrete-intuitive KE, 'current is nonlinear/wavelike', while High and Medium groups shared strong connections to canonical KE, 'current reaches peaks and lows repeatedly'. Medium group also made strong connections to K2, 'observed and calculated time constants of current match' (Figure 8.4 UCG, Appendix 8). For PostQ2, Low group was found to make most connections to K2, 'V, I behavior similar', while Medium and High groups shared common connections to K1, 'voltage, current increase-decrease proportionally', and to K2, 'V, I have inverse relation'. Medium group also made less strong connections to concrete-intuitive KE, 'V, I have direct relation' (Figure 8.4 URG, Appendix 8).

Test1. At an alpha-level of .05, there were no significant differences found between mean epistemic networks of Low, Medium and High score groups of Test 1, for PreQ, PostQ1, and PostQ2

(Table 8.5, Appendix 8). For PreQ, Low and High score groups were found to make strong connections to canonical KEs, ‘time to reach maximum value will increase’, and ‘current/circuit will on-off at intervals’. Medium and High groups shared connections to canonical KEs, ‘current will stabilize when capacitor charges’, and ‘current will decrease’, and to abstract-intuitive KE, ‘current will increase’ (Figure 8.5 Upper left graph, Appendix 8). After the experiment, Low and High groups shared strong connections to concrete-intuitive KE, ‘current is nonlinear/wavelike’, while medium and high groups shared connections to K2, ‘observed and calculated time constants of current match’. High group was also found to make strong connections to K1, ‘current reaches peaks and lows repeatedly’. (Figure 8.5 UCG, Appendix 8). For PostQ2, High and Low groups were found to share strong connections to K1, ‘voltage, current increase-decrease proportionally’, while Medium group made most connections to concrete-intuitive KEs, that voltage and current have ‘direct’, and/ ‘inverse’ relation. All three groups also shared less strong connections to K1, ‘V, I have same time constants’, and to K2, ‘V, I behavior similar’ (Figure 8.5 URG, Appendix 8).

Summary of Network Analysis Results

The ENA Web Tool modelled connections (edges) between E, K1, K2, and K3 (nodes), into adjacency matrices, by the number of times the codes or knowledge elements co-occurred. Adjacency Matrices were then normalized, rotated, and converted by the Tool, into individual and/mean epistemic networks (graphs), based on the type of inputs or variables (units, stanzas, and conversations) chosen. These were chosen based on the analyses that I wanted to conduct and/ explore (Chapter 3). The structure of connections in each network graph - the thickness of edges, their lengths, node thickness (weight), and the relative positions of nodes in the network were derived automatically by the Tool, using complex mathematical formulae, based on the number of times each connection and node co-occurred in student response, to enable maximum variance, necessary for optimal visualization of networks. Since the co-occurrences changed from PreQ to PostQ, and the nodes changed from experiment to experiment, the network structure or the pattern of connections was also found to change from one experiment to another, from PreQ to PostQ, and by the type of analyses conducted.

However, four broad categories of connections (edges), formed by four broad themes (nodes) identified through qualitative analyses, were found to emerge from the various network analyses conducted: (i) canonical connections, or connections between epistemic construct and canonical knowledge elements (E-K1, and K1-K1); (ii) transitional connections, or connections between canonical, concrete-intuitive, and abstract-intuitive knowledge elements (K1-K2, K1-K3); (iii) intuitive type 1 connections, or connections between epistemic construct, and concrete-intuitive knowledge elements (E-K2, K2-K2); and (iv) intuitive type 2 connections, or connections between epistemic construct, abstract-

intuitive, and concrete-intuitive knowledge elements (E-K2, K3-K3, and K2-K3). The presence or absence of specific nodes and edges, and their weights, and strengths of association, respectively, determined the structure of each network, and it was found to vary by experiment and analysis.

Experiment 1. Before the experiment, strong concrete-intuitive connections were found to K2, ‘experience shock’, less strong canonical connections to K1, ‘wet hands better conductor’, and less strong transitional connections between this K1 and K2, were also found in the mean network. After experiment, mean network shifted to strong canonical connections to K1, ‘wetness decreases hand resistance’, along with weak connections to other K1, K2, and K3 KEs. There were no statistically significant differences found between mean networks of low, medium and high score groups of EBAPS, lab Quiz, Labs, and Test1, and between the two sections of the course. Significant differences were found between mean networks of low and high scorers on A3 homework quiz, with the low scorers having less diverse networks, but strong CC, TC, and I1C, and high scorers being different because of more number of weak I2C, along with strong CC, TC, and I1C, like low scorers. After the experiment, low and medium groups differed, because of the same reasons. Thus, networks of low scorers were found to be more precise (having limited, but strong CC, TC and I1C) and less diverse.

Experiment 2. Significant differences were found between mean networks before and after experiment 2 (PostQ). Although strong canonical connection to K1, ‘change distance’, and less strong, to K1, ‘change area’, were found in both mean networks, strong concrete-intuitive connections to K2, ‘change number of pages’, were also found in mean network for PostQ1. For the interpretive question (PostQ2), strong canonical connections were found to K1, ‘capacitance decreases with increase in distance’, while canonical connections to ‘change in area’ were weak. The PostQ2 network had more diverse but weak, TC, I1C, and I2C connections. Significant differences were also found between mean epistemic networks of students from sections 1 and 3, both before and after the experiment, with section 3 network showing strong CC to ‘change area’ and ‘change distance’ unlike section 1 network, which showed strong CC only to ‘change distance’, before the experiment. Post-experiment network of section 3, showed stronger I1C to ‘change number of pages’, than section 1. For PostQ2, section1 network had stronger CC to ‘area’ and ‘distance’ relationships of capacitance, and section 3 network had more TC between ‘area’ and ‘distance’ relationships, which meant that these students were relating to both physical quantities as being significant for changing capacitance.

There were no statistically significant differences found between the mean networks of high, medium and low score groups for EBAPS, A3 homework quiz, labs, and test1. High and low score groups of Lab Quiz were found to have significant differences in their networks before experiment, with high scorers making more CC and I1C to ‘change in area’, and ‘change in overlap area’.

Experiment 3. Significant differences were found between mean epistemic networks of PreQ2, and PostQ1, PostQ2. This is because, Post-Q networks had strong CC to ‘nonlinear graph/non-constant slope’, while PreQ network had strong CC to diverse K1s, such as ‘do not follow ohm’s law’, and ‘have nonlinear VI relation’, along with graphical K1 found in Post-Q networks. Other weak CC, TC, I1C and I2C connections of PreQ and PostQ, were also found to be different, in that the PostQ connections were to ‘nodes’ referring to linearity or non-linearity of graphs. Significant differences were also found between mean networks of sections 1 and 3 for PreQ2, and post question pertaining to light bulb (PostQ1). Section3 networks showed strong CC to ‘nonlinear graph/non-constant slope’, both before the experiment, and after the ‘light bulb’ activity. No significant differences were found between low, medium, and high score groups of EBAPS, for pre- and post- experiment questions.

For PreQ2, networks of Low and Medium score groups of A3 HW Quiz, and of Test1; Low and High score groups of Lab Quiz, and of Lab4; and Medium and High score groups of A3 HW Quiz, were found to be significantly different. Networks of Low and high scorers of A3 Quiz, had strong CC to graphical KEs, as compared to medium scorers. Low scorers of Lab Quiz, and Lab4, were found to make fewer CC to graphical KEs, as compared to corresponding high score groups. For the light bulb question, Low and Medium score groups, and Low and High score groups of A3 HW Quiz, and Lab4; and, Low and High score groups, and Medium and High score groups of Lab Quiz, were found to have significantly different networks. The striking differences seemed to be, that high scorers made more CC to graphical KEs, and low scorers made less abstract-intuitive or I2 connections. For the LED question, Medium and High score groups of Lab Quiz and Test1, and Low and High score groups of Lab Quiz, were found to have significantly different networks. On an average, Medium score group made diverse connections to K1, K2, K3; low scorers were found to connect to concrete-intuitive KEs, and high scorers made strong CC to graphical KEs.

Experiment 4. Significant differences were found between learners’ mean epistemic networks for PreQ and PostQ1. Pre-experiment networks showed strong canonical connections to K1s, ‘current will decrease’, ‘will stabilize when capacitor charges’, and/ ‘time to reach maximum value will increase’, while post-experiment networks demonstrated different K1, ‘current reaches peaks and lows repeatedly’, and significantly more I1 connections, to K2s, ‘current is nonlinear or wavelike’, and/ ‘experimental and calculated time constants of current are same’. Post Q2 networks, also showed stronger I1C connections than for PreQ. There were no statistically significant differences found, between mean networks of section 1, and section 3 students, and between mean networks of low, medium, and high score groups of Test1, for all questions.

Mean networks of Low and Medium score groups of A3 HW Quiz differed significantly for PreQ and PostQ2. Before experiment, Low and Medium groups made strong CC to different K1s, and after the experiment, Low group made strong connections to concrete-intuitive KE, while Medium group made strong CC. *For PreQ*, Low and High score groups of Lab Quiz, and Medium and High score groups of EBAPS, were found to differ significantly, since they made strong CCs to different K1s. Low group connected strongly to ‘current will stabilize when capacitor charges’, Medium group connected to, ‘circuit will on-off at regular intervals’, while High group connected to, ‘time to reach maximum value will increase’. All 3 groups shared common connections to K1, ‘current will decrease’. *For PostQ1*, Low and Medium score groups of EBAPS, and Low and High score groups of Labs, were found to have significantly different mean networks. Medium and High score groups were found to make strong CC to K1, ‘current reaches peaks and lows repeatedly’, while Low score groups made strong I1 connections, to K2, ‘current is nonlinear/behaves wavelike’.

To sum up, Pre-and Post- experiment mean network structures differed significantly for all four experiments. Differences in network structures were also found between Sections of course for Experiments 2 and 3. Existence of significant differences between Mean networks of Low, Medium and High score groups, varied by experiment and question, as discussed above. Differentiation of networks by scores, was found for A3HW Quiz, on all experiments except experiment 2; for EBAPS on Experiment 4; for Lab Quiz on Experiments 3 and 4; for Labs on all experiments except experiment 1; and for Test 1, on experiment 3. Thus, structural differences were found to be more sensitive to A3 HW Quiz, Lab Quiz, and Lab4. The following section will discuss the findings of quantitative analyses conducted using RStudio, on the same rData, which was plotted as network graphs, and analyzed using network analysis techniques, as discussed in this section.

Quantitative results

This section will report the computation of measures of sophistication, descriptive quantitative results for each of these variables, and results of correlational analyses conducted to test the validity and reliability of these variables, to answer RQ3: (a) What levels of sophistication were identified in learners’ thinking, through exploratory statistical analyses, of scores of node-centrality and weight density of connections? (b) How do the measures of centrality and connection weights, of learners’ Epistemic Networks relate to (i) Lab Score (ii) Quiz scores, (iii) Test1score, (iv) EBAPS score, and (v) A3 Home Work Quiz score? (QUAN).

Each of the twelve rData files (one per question), exported to RStudio from the ENA WebTool, were processed by running an R-script developed by me (Appendix 9), to extract data required to compute the measures of centrality and weight density, from each ENA dataset. The ‘output table’

exported to Spreadsheet from RStudio, contained (i) 18 columns of metadata - quantitative test scores of EBAPS, A3HWQuiz, Lab Quiz, Lab, and Test 1, Section of course, Gender, Student ID, group membership by quantitative scores, (ii) 1 column of raw textual data, (iii) code columns containing values “0” or “1”, depending on presence or absence of a code in student’s response, (iv) 6 columns of centroid positions of each student’s epistemic network in the network projection space along six dimensions – svd1, svd2, svd3, svd4, svd5, and svd6, and (v) connection weight columns for all possible connections, for a given set of codes (Figures 3.20 to 3.22 in Chapter 3). Table 4.19 below presents a summary of total number of possible connections for each experiment and question.

Table 4.19 Node(code) and Edge (connection) Summary by Question and Experiment

	Expt1 Codes	Connections	Expt2 Codes	Connections	Expt3 Codes	Connections	Expt4 Codes	Connections
PreQ1	8	28	7	21	9	36	10	45
PreQ2					9	36		
PostQ1	7	21	7	21	9	36	7	21
PostQ2			13	78	10	45	6	15

The connections were classified into four different categories based on the types of nodes (code themes) which connected to each other, to make up a connection (Table 4.20).

(i) Canonical connections (CC) were defined as connections between canonical knowledge elements (K1), and the epistemic construct (E), or between two canonical KEs;

(ii) Transitional connections (TC) were defined as connections between canonical knowledge element and abstract-intuitive, or concrete-intuitive knowledge element, signifying a transition from intuitive to canonical concept formation;

(iii) Concrete-intuitive or Intuitive 1 connections (I1C) were defined as connections between epistemic construct, and concrete-intuitive knowledge elements (K2), or between two concrete-intuitive KEs;

(iv) Abstract-intuitive or Intuitive 2 connections (I2C) were defined as connections between epistemic construct, and abstract-intuitive knowledge elements (K3), or between two abstract-intuitive KEs, or between concrete-intuitive and abstract-intuitive KEs.

Table 4.20 Types of Connections

Canonical (CC)	Transitional (TC)	Intuitive 1(I1C)	Intuitive 2 (I2C)
E – K1	K1 - K2	E - K2	E – K3
K1- K1	K1 – K3	K2 – K2	K2 – K3
			K3 – K3

The cell values in all connection columns of the output table, obtained from RData generated by the ENA Web Tool, equal the weight of the connection, or the strength of association. These values were used to compute the measures of ‘connection weight’ and of ‘node centrality’, as per their definitions specified in the ENA Tutorial and User Manual documents, and in consultation with the researchers from University of Wisconsin, who developed the ENA WebTool.

Measures of Sophistication of Thinking

The computation of sophistication measures, consisted of seven Excel function commands presented in Appendix 10. Each function computed one out of seven variables (assessment scores), created in seven new columns, as follows:

1. K1- centrality, or centrality of K1 knowledge elements in learners’ thinking: For each row, K1C was computed by taking the sum of connection weights of all cells, whose column (connection) contained K1.
2. K2 – centrality, or centrality of K2 knowledge elements in learners’ thinking: For each row, K2C was computed by taking the sum of connection weights of all cells, whose column (connection) contained K2.
3. K3 – centrality, or centrality of K3 knowledge elements in learners’ thinking: For each row, K3C was computed by taking the sum of connection weights of all cells, whose column (connection) contained K3.
4. Canonical connection strength: For each row, CCS was computed by taking the sum of connection weights of all cells, whose columns contained either E – K1, and/ K1- K1 connections.
5. Transitional connection strength: For each row, TCS was computed by taking the sum of connection weights of all cells, whose columns contained either K1 – K2, and/ K1- K3 connections.
6. Intuitive 1 or concrete-intuitive connection strength: For each row, I1CS was computed by taking the sum of connection weights of all cells, whose columns contained either E – K2, and/ K2- K2 connections.

7. Intuitive 2 or abstract-intuitive connection strength: For each row, CCS was computed by taking the sum of connection weights of all cells, whose columns contained either E – K3, K2 – K3, and/ K3- K3 connections.

Levels of sophistication (S-Level) were then developed using connection strengths (weights), since node-centrality measures are essentially subsumed within the four measures of connection weight density, by design, as was also corroborated by the results of descriptive statistics presented below. Combinations of co-occurring connections present in each dataset were examined by sorting the data by connection weights, and tabulated for each question, and experiment (Appendix 11). Then criteria were identified, and five levels of sophistication defined (Table 4.21).

Table 4.21 Criteria to Assess Level of Sophistication of Thinking

Criterion	S-Level
If I2C occurs alone at any value of connection strength	1
If I2C co-occurs with any other connection, and has greatest strength	1
If I1C occurs alone at any value of connection strength	2
If I1C co-occurs with any other connection, and has greatest value of strength	2
If TC occurs alone at any value of strength	3
If TC co-occurs with any other connection, and has greatest value of strength	3
If CC co-occurs equally with other connections, but strength less than 1	3
If CC occurs alone and is equal to or less than 1	4
If CC co-occurs with any other connection, and has greatest strength ≤ 1	4
If CC occurs alone and has strength greater than 1	5
If CC co-occurs with TCS and has greater strength which is also greater than 1	5

Sophistication measures and S-Level data for each question, and experiment, were analyzed with the help of SPSS, to obtain descriptive statistics - mean, median, and percent count of students who scored each observed value of score, to answer RQ3(a). Correlational analyses were conducted between student scores on five quantitative assessments used in the course - EBAPS, A3HW Quiz, Lab Quiz, Lab score, and Test 1, to answer RQ3(b). Findings of these analyses are presented below, for each experiment.

Descriptive Statistics: Experiment 1

Table 12.1 (Appendix 12) presents a summary of network measures of interest in the context of this study – node-centrality, and edge weight (connection strength). Figure 4.19 presents a summary of sophistication levels of students for Experiment 1, PreQ and Post Q. Key findings are discussed below.

Connection (edge) weight. After the experiment, median Canonical connection strength (CCS) increased from 0.5773 to 1; percentage of students with high score (maximum) increased from 2.6 to 28.6,

and percentage with low score decreased from 42.1 to 11.9. While 55.3% of students did make Transitional connections having strength of at least 0.5774, before the experiment, 97% of students were found to make no transitional connections (K1-K2, and/K1-K3), after the experiment. However, although the percentage of students with abstract-intuitive connections decreased from 10.6 to 4.8, and the values of I2CS were also found to decrease, 20% of the students retained their concrete-intuitive connections after the experiment. These students would benefit from ‘explicit directions’ by the TAs, to help them resolve issue with their thinking, as discussed at length in previous sections (2 and 3).

Node centrality (Node weight). Centrality of canonical KEs, K1C, was found to remain unchanged in learners’ thinking after the experiment, since 57% of students made almost the same number of connections to K1, before and after the experiment. However, centrality of K2, decreased, from 80% to 5%, while that of abstract-intuitive elements, remained almost unchanged at 5%. These findings illustrate, that connection weights afford better differentiation, than node weights. This is to be expected, because node weights sum over all possible connections to K1, essentially equivalent to code counts. Hence, connection weights have been chosen to differentiate sophistication levels.

Sophistication level. Figures 4.19 shows that the median sophistication level of participants increased from level 3, to level 4 after the experiment. Students demonstrated levels 1, 2, and 3, of sophistication, in increasing order of percent count, before experiment, and levels 4 and 5, in decreasing order of percent count, after the experiment. Thus, the experiment was effective in enhancing connections in learners’ thinking, from a dominance of transitional connections, to canonical connections.

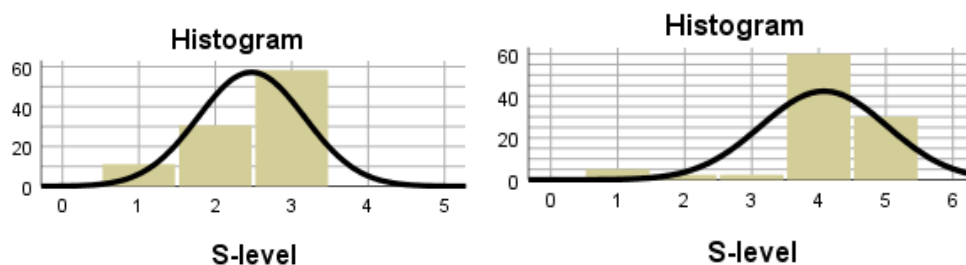


Figure 4.19 . Frequency distribution of levels of sophistication for PreQ (left chart) and Post Q (right chart).

Descriptive Statistics: Experiment 2

Table 12.2 (Appendix 12), presents a summary of node-centrality, and edge weights (connection strength). Figure 4.20 presents a summary of sophistication levels of students, for Experiment 2, PreQ, Post Q1, and PostQ2. Key findings are discussed below.

Connection weight. After the experiment (PostQ1), median CCS score was found to decrease from 1 to 0.5773, and percentage of students who made no canonical connections at all, increased from 20 to 46. However, for the question asking them to interpret the relationships between capacitance, area, and distance (PostQ2), the median values of canonical connections, and the overall percentage of students making canonical connections of strength zero, less than 1, 1, or greater than 1, remained the same as for PreQ. While the median strength of Transitional connections remained the same before and after experiment, percentage of students making Transitional connections was found to decrease after the experiment, for both PostQ1 and PostQ2. On the other hand, the median strength of concrete- intuitive connections was found to increase from 0 to 0.5773, and the percentage of students making these connections increased from 13, before the experiment, to 54, after the experiment, for PostQ1.

For the interpretation question, no changes were found in the median score and the percent count, for concrete-intuitive connections. With respect to abstract-intuitive connections, the percent count of students was found to decrease from 40 to 5, for PostQ1, and from 40 to 27, for PostQ2.

Node centrality (Node weight). Centrality of canonical KEs, K1C, was found to decrease in learners' thinking after the experiment (PostQ1), since percent count of students with K1 centrality weight of 1 or greater than 1, decreased from 77% to 42 %. However, percent count of students reflecting centrality of canonical KEs in their responses to Post Q2, remained the same as for PreQ. Percent count of students demonstrating a weight of 1 or more, for abstract-intuitive knowledge elements, decreased from 35% to 5% for PostQ1, and to zero, for PostQ2. At the same time, percent count of students who demonstrated centrality of concrete-intuitive knowledge elements in their thinking, with weight of 1 or more, increased from 13% before the experiment, to 52% after the experiment for PostQ1; but did not change for PostQ2.

Sophistication level. Results for PostQ2, which required students to make interpretations of relationships governing the change of capacitance, unlike PostQ1, which asked them to describe what they did to change capacitance, are being considered to assess sophistication, for this experiment. Results for PostQ1, showed a decrease in overall levels of sophistication, possibly because the structure of PostQ1, directed students' attention towards 'concrete' entities, thereby influencing their responses. By analyzing histograms for PreQ, and PostQ2 (Figure 4.20) it was found, that that the percent count of students who demonstrated high level of sophistication (level 5), making canonical connections with a strength greater than 1, increased by 25%, while for level 4 (CCS = 1), it decreased by 18% after the experiment.

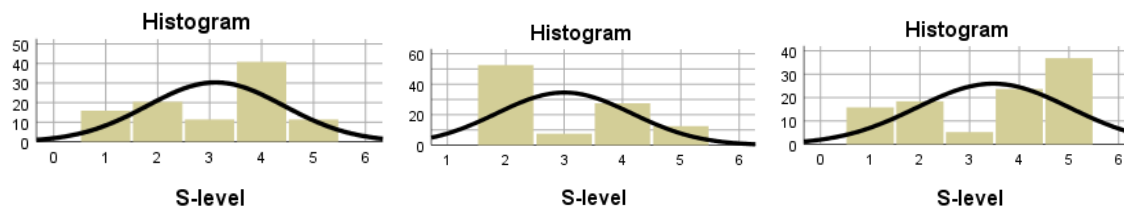


Figure 4.20. Frequency distribution of levels of sophistication for PreQ (left), Post Q1 (center), and PostQ2 (right).

Thus, the ‘sophistication of thinking’ was enhanced by the experiment, for at least 23% of students from level 3 or 4, to 5, and the percent count of students with level 1, or level 2, of sophistication, was not found to change after the experiment. However, it also means, that for 35% of students with S-levels of 1 or 2 before the experiment, the experiment was not effective in redirecting, scaffolding, or reorienting their frames, and these students could benefit from explicit, contextual directives from the TAs.

Descriptive Statistics: Experiment 3

Table 12.3 (Appendix 12) presents a summary of node-centrality, and edge weights (connection strength). Figure 4.21, presents a summary of sophistication levels of students, for Experiment 3, PreQ2, Post Q1, and PostQ2. Key findings are discussed below.

Connection weight. After the experiment, percentage of students who made no canonical connections at all, decreased from 46 to 23 for LED (PostQ2) question, and increased by 2% for bulb (PostQ1) question. Consequently, percentage of students making canonical connections of strength greater than 1, was found to increase by 26% for PostQ2, while it decreased by 5% for PostQ1. Percentage increase of Transitional connections was 5 and 8 for bulb and LED questions respectively. The maximum strength of abstract-intuitive connections was found to be greater than 2, in the case of LED, and the percent counts for I2C and I1C, increased by 17%, and 8% respectively, after the experiment. The percent count increase of intuitive connections was less for the bulb question, than it was for the LED question – 5% for I1C, and 2% for I2C.

Node centrality (Node weight). Centrality of canonical KEs, K1C, was found to increase in learners’ thinking after the experiment, for LED, by 20%, while it remained unchanged for bulb. Percent count of centrality of concrete-intuitive (K2C), and abstract-intuitive (K3C) knowledge elements, increased by 8% and 18% respectively, for LED, while it increased by 10% each, for the bulb. Thus, node-centrality, while being indicative of trends in thinking, does not provide a nuanced picture of the structure of learners’ epistemic networks, unlike edge weights.

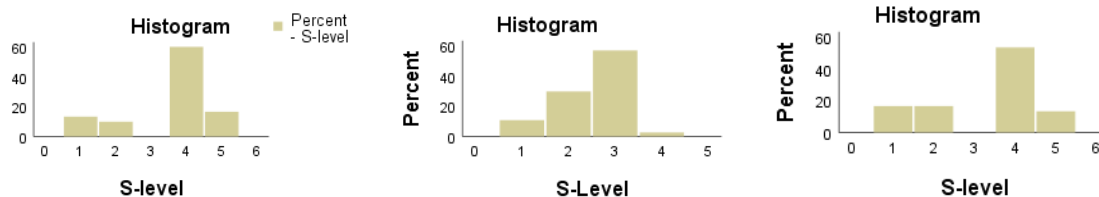


Figure 4.21. Frequency distribution of levels of sophistication for PreQ2 (left), Post Q1 (center), and PostQ2 (right).

Sophistication level. Of all respondents to PreQ2, which asked students to describe non-ohmic materials, approximately 80% responses were canonically appropriate, at a sophistication level of 4 and 5, while approximately 20% responses were intuitive, at sophistication levels of 1 and 2. The post questions asked them to classify bulb, and LED as ohmic or non-ohmic, based on their findings from the experiment, and to justify their choices. For the light bulb question, the percentage of students who demonstrated higher levels of sophistication, (levels 4 and 5) by using canonical language, decreased from 80% before the experiment, to less than 5%, since most of these students transitioned to levels 3, 2 and 1. Thus 60% of students were found to make transitional connections (level 3), while the percentage of students making intuitive connections was found to increase from 20% before experiment, to 40% after experiment, for the light bulb question. Their classification of bulb, canonically, as a non-ohmic material, was hindered by their prior unresolved ideas about the classification of graphs as linear and non-linear (as discussed in Sections 2 and 3 of this Chapter).

For the LED question, percentage of students thinking in terms of intuitive knowledge elements (levels 1 and 2), also increased by about 10% after the experiment. There was a corresponding decrease in percentage of students demonstrating higher levels of sophistication in terms of using canonical language to classify, and justify LED as non-ohmic, from 80% to 70%. Thus, either, the students sought more help from TAs, to understand the graphical output of LED, and therefore, were largely able to justify LED, correctly, since they were clearly unable to interpret the graphs correctly in case of light bulb; or their preconceived intuitive notions of the bulb being ohmic, interfered with their classification of bulb as non-ohmic, rather than their intuitive interpretations of graph. Either way, such epistemological ‘hindrances’ need to be addressed in-context, in the broader interest of enhancing access to physics; and mixed methods interpretations of qualitative information on the substance and process of learners’ thinking, can provide a mechanism for TAs and instructors to monitor and address the missing, missed, and/or misconceptions in learners’ thinking.

Descriptive Statistics: Experiment 4

Table 12.4 (Appendix 12) presents a summary of node-centrality, and edge weights (connection strength), and Figure 4.41, presents summary of sophistication levels of students, for Experiment 4, PreQ, Post Q1, and PostQ2. Key findings are discussed below.

Connection weight. After the experiment, percentage of students who made no canonical connections at all, increased from 36 to 64 for PostQ1, and to 67 for PostQ2. Consequently, percentage of students making canonical connections of strength equal to, or greater than 1, was found to decrease from 62% before the experiment, to 32, and to 33, for PostQ1, and PostQ2, respectively. At the same time, percentage of students making Transitional and concrete-intuitive connections, increased by 3%, and 13%, for PostQ1, while it decreased by 4% for abstract-intuitive connections. For PostQ2, which required students to compare the behaviors of voltage and current in RC circuits, unlike PostQ1, which asked them to describe findings the behavior of current, based on experimental findings, percentage of students making transitional and concrete-intuitive connections, was found to increase by 9%, and 19% respectively, while percent count of students making abstract-intuitive connections decreased by 13%. This could mean, that the experiment was effective in providing concrete anchors to 13% of students, and transitioning their thinking from abstract-intuitive knowledge elements, to concrete-intuitive knowledge elements. However, at least 15% of students who demonstrated canonical thinking before the experiment, were found to transition to concrete-intuitive connections, possibly, because PostQ1, explicitly drew their attention towards findings.

Node centrality (node weight). Centrality of canonical knowledge elements in learners' thinking, decreased by 30%, after the experiment, for both PostQ1 and PostQ2. Centrality of abstract-intuitive KEs, also decreased by 7% and 16% after the experiment, for PostQ1, and PostQ2, respectively. Correspondingly, it was found, that the centrality of concrete-intuitive elements increased by 13% for PostQ1, and by 26% for PostQ2. Thus, after the experiment, learners' responses were influenced by the findings of the experiment, and therefore, the centrality of concrete-intuitive elements increased, even for PostQ2, which required students to make interpretations about similarities and differences between the behaviors of voltage and current, in RC circuits.

Sophistication level. By analyzing the 3 charts in Figure 4.22 below, it was found that the percentage of students whose responses reflected levels 1 (abstract-intuitive connections) and 2 (concrete-intuitive connections) of sophistication of thinking, increased by about 15% after the experiment, for PostQ1, and by about 8%, for PostQ2. Percentage of students whose responses demonstrated sophistication level 4, before the experiment, decreased approximately, by 13% for PostQ1, and by 20%, for PostQ2, while percent count of students demonstrating sophistication level 3 of thinking, increased by about 5% and 15%, for PostQ1, and PostQ2, respectively. Thus, for this experiment, learners did not

demonstrate sophistication level above 4, and their thinking reflected more transitional connections after the experiment, than before, suggesting that the experiment was not very effective in enabling strong canonical connections in learners' thinking.

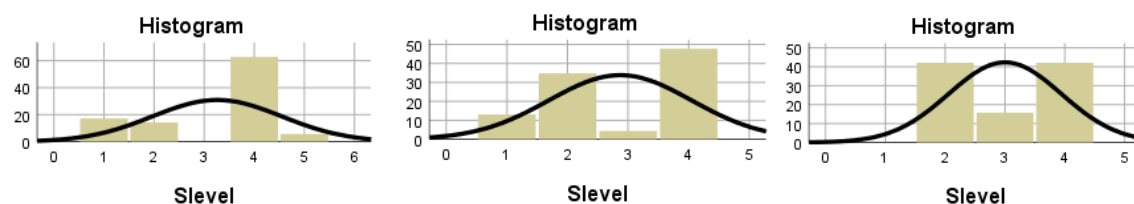


Figure 4.22. Percent distribution of levels of sophistication for PreQ (left), Post Q1 (center), and PostQ2 (right).

Correlational Analyses

This subsection presents the findings of correlation analyses conducted using SPSS, between the seven measures of sophistication – CCS, TCS, I1CS, I2CS, K1C, K2C, K3C and learners' scores on four quantitative instruments used in the course – A3 Homework Quiz, Lab Quiz, Lab4 scores, Test 1, and EBAPS, by experiment and question. Spearman's rho non-parametric correlation was chosen for the analyses, since quantitative scores were not normally distributed (as found in Section1). Significant values of correlation coefficients are marked with single asterisk for significance level of 0.05, and with double asterisk for significance level of 0.001, in the Tables of findings presented in Appendix 13. Theoretically, negative correlations were expected between measures corresponding to lower levels of sophistication - K2C, K3C, I1CS, I2CS, and existing assessment measures – A3Quiz, Lab Quiz, Labs, and Test 1. Positive correlations were expected between measures corresponding to higher levels of sophistication – K1C, and CCS. TCS, being a transitional measure (at level 3 of sophistication), was expected to correlate either positively or negatively, with existing assessments.

Experiment 1. Table 13.1 (Appendix13) presents the results of correlational analyses of Experiment 1. For PreQ, no significant correlation was found between sophistication scores assessed by ENA, and scores of regular assessments. For PostQ, weak negative correlation was found between EBAPS scores and I1CS, and K2C scores. Thus, learners with high scores on the Epistemological Beliefs About Physical Sciences survey, were found to have low scores on centrality of concrete-intuitive knowledge elements, and connection strength of concrete-intuitive connections in their thinking. This finding could be viewed as evidence of concurrent validity between EBAPS, and I1CS, and K2C. Weak positive correlation was also found between Test1 scores, and K1C, which agrees with the expectation

that higher centrality of canonical knowledge elements could lead to higher scores on Test1. This finding could be counted as evidence in favor of convergence validity between Test 1 and K1C.

Experiment 2. Table 13.2 (Appendix13) presents the results of correlational analyses of Experiment 2. For PreQ, weak negative correlation between Lab score and strength of transitional connections was found, meaning students with lower scores on strength of transitional connections were found to have higher Lab scores, which agrees with theoretical expectations about TCS. However, medium positive correlation was also found between Test 1 scores, and strength of transitional connections, for PostQ1, confirming the hypothesis that TCS is a transitional measure, and therefore, not a valid criterion of sophistication. For PostQ2, medium positive correlation was found between Lab4 scores, and K2 centrality, and I1CS scores. Although this finding is contrary to expectation, the reason for this discrepancy, could be that PostQ2 responses had close correspondence with students' Lab4 responses, and therefore, the correlation found. Weak correlation was also found between learners' A3 HW Quiz scores and canonical connection strength, as expected, and could be considered indicative of convergent validity between A3Quiz, and CCS.

Experiment 3. Table 13.3 (Appendix13) presents the results of correlational analyses of Experiment 3. Since TCS scores were zero for PreQ2 and PostQ2, corresponding values of correlation coefficients, show as 'missing' in the Table. For PreQ2, no correlations were found between sophistication scores and existing quantitative assessment scores. For PostQ1, weak negative correlations were found between Lab Quiz, Lab4 scores, and K2C and I1CS, meaning that students with lower scores on centrality of concrete-intuitive knowledge elements, and on strength of concrete-intuitive connections, were found to have higher scores on Lab4, and Lab Quiz. This finding is also in agreement with the hypothesis, and counts as evidence of convergent validity between Lab4, Lab Quiz, and K2C, and I1CS. Weak negative correlation was found between A3 HW Quiz and strength of transitional connections, meaning that students with higher values of A3Quiz scores, were found to have weak transitional connections, as could be expected. For PostQ2, weak positive correlations were found between EBAPS and I1C, meaning that students with higher EBAPS scores were found to have higher scores on concrete-intuitive connection strength. This finding is however, contrary to expectation, and could be explained by the fact that Post-experiment responses for Experiment 3, were mostly based on findings of the experiment, and therefore, included more references to concrete-intuitive knowledge elements.

Experiment 4. Table 13.4 (Appendix13) presents the results of correlational analyses of Experiment 4. For PreQ, and PostQ2, no correlations were found between sophistication scores and scores on existing quantitative assessments. For Lab4, positive correlation was found between lab scores and centrality of abstract-intuitive knowledge elements and abstract intuitive connection strength, which

seems to suggest that even students with high lab scores, made significant connections to abstract-intuitive knowledge elements, for Experiment 4. This finding is also, therefore, contrary to expectation.

Summary of Quantitative Findings

Five levels of sophistication could be identified in learners' thinking, by computing three measures of node-weight (centrality), and four measures of edge weight (connection strength), from the ENA datasets generated by ENA Web Tool, for each question and experiment, and by analyzing these seven variables, using Descriptive statistics. Level 1 of sophistication was characterized by novice-like epistemic networks, having significantly strong Intuitive 2 (abstract-intuitive) connections; level 2 by epistemic networks having strong Intuitive1 (concrete intuitive) connections; level 3 by strong transitional connections, level 4 by strong canonical connections, and level 5, by very strong canonical connections. Epistemic gains, or gains in levels of sophistication differed by experiment (Table 4.22 below).

Table 4.22 Summary View of S-level Gains by Experiment

	Expt. 1		Expt. 2		Expt. 3		Expt. 4	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
PreQ	2.47	3	3.11	4	3.57	4	3.26	4
PostQ1	4.08	4	3	2	2.51	3	2.87	3
PostQ2			3.47	4	3.47	4	3	3

For experiment 1, the median sophistication level of participants increased from level 3, to level 4 after the experiment. For experiments 2 and 3, median sophistication levels did not change after the experiment for PostQ2, which was an interpretive question for Experiment 2, and the LED question for Experiment 3. However, the median sophistication levels were found to drop by one level for the light bulb question of Experiment 3, and by 2 levels, for the descriptive question of Experiment 2. For experiment 4, there was a drop in median sophistication level for both questions. Thus, positive sophistication gains were realized only by Experiment 1. The remaining three experiments yielded net negative gains with respect to sophistication of thinking.

Correlational analyses between sophistication scores on ENA-based assessments, and scores on existing quantitative assessments, revealed mixed results. Table 4.23 presents a summary of findings.

Table 4.23 Correlation Summary by Experiment and Question

	Positive Correlations	Question	Negative Correlation	Question
Experiment 1	Test 1 – K1C	PostQ	EBAPS – K2C EBAPS – I1CS	PostQ PostQ
Experiment 2	Test 1 - TCS A3Quiz - CCS Lab 4 – K2C Lab4 – I1CS	PostQ1 PostQ2 PostQ2 PostQ2	Lab 4 - TCS	PreQ
Experiment 3	EBAPS – I1CS	PostQ2	A3Quiz - TCS Lab Quiz – K2C Lab Quiz – I1CS Lab 4 – K2C Lab 4 – I1CS	PostQ1 PostQ1 PostQ1 PostQ1 PostQ1
Experiment 4	Lab 4 – I2CS	PostQ1		

While results of correlational analyses aligned well with theoretical expectation, they can only be taken as indicators of correlation, rather than conclusive evidence of convergent and/ concurrent validity, because (a) the absence of correlation or weak correlations could also be construed as evidence of discriminant validity, or the fact that the constructs of sophistication, and the constructs assessed by Quizzes, labs, and/ tests are not the same; (b) the qualitative coding was not aligned with the rubrics used to assess responses of existing assessments, since the goal was to test the viability of using a technology interface for assessment of thinking, rather than to develop a criterion referenced assessment; and (c) the experiments differ significantly in terms of difficulty level, and prior knowledge required, meaning correlations found in one experiment, cannot be extrapolated to other experiments.

Mixed Methods Analyses: Data Integration

This section will integrate the findings and results of qualitative and quantitative analyses, which were reported in previous sections, to answer RQ4: (a) How do the inquiry activities compare with respect to their effect on the sophistication of learners’ thinking? (b) Does the proposed ENA-based assessment framework provide a valid and reliable tool to assess the sophistication of students’ thinking in inquiry contexts in undergraduate physics courses?

Table 4.24 below, presents a synopsis of the analyses conducted in this study, which includes the purpose of conducting these analyses, research question targeted if any, and the outputs (results) generated.

Table 4.24 Synopsis of Data Analyses and Findings Presented in This Chapter

Step	Data (Input)	Type	Analysis	Type	Purpose	RQ	Results (Output)
1	EBAPS	QUAN	Scoring with Template Rubric	QUAN	Determine Prior S- level		EBAPS Scores
2	A3Quiz, Lab Quiz, Labs, EBAPS, Test 1	QUAN	Descriptive Statistics; percentiles	QUAN	Grouping by scores; Validity Testing		Low, Medium, High Score groups
3	Twitter data	QUAL	Thematic Coding	QUAL	Identify Knowledge themes	RQ1	Epistemic Construct, Canonical Physics, Concrete-intuitive physics, Abstract-intuitive physics
4A	Coded Twitter data	MIXED	Network Analyses	MIXED	Identify Connection Pattern	RQ2	Canonical connections, Transitional connections, Intuitive 1 connections, Intuitive 2 connections
5	Coded Twitter data	MIXED	Network Analyses; t-tests	MIXED	Discriminant Validity	RQ2	Group Differences in epistemic networks
6	rdata	QUAN	Descriptive Statistics	QUAN	Compute S-scores	RQ3	CCS, TCS, I1CS, I2Cs, K1C, K2C, K3C, S-level
7	S-Scores	QUAN	Correlational Analyses	QUAN	Validity and Reliability	RQ3	More work needed for validity and reliability
8B	All Results	MIXED	Data Integration	MIXED	Data Integration	RQ4a	S-gains positive only for Experiment 1
9C	Correlation Results	MIXED	Data Integration	MIXED	Validity and Reliability	RQ4b	Discriminant Validity Internal Consistency Reliability

The purpose of data integration in this study, was to provide complementarity – to elaborate, enhance, illustrate and classify, results, and to enable development of one analyses from the results of another. So, data integration was done at three points in the study, labeled A, B, and C, in Table 4.24. At (4) A, the grouping results from Step 2, the input data of step 2, and the EBAPS scores generated from Step 1 – the metadata, was integrated with the input textual data of Step 3, and the code column results generated in step 3. The integrated data was then loaded into the ENA WebTool. In this section I will draw upon the results of previous sections and integrate the findings to answer RQ4a (integration point 8B), and RQ4b (integration point 9C). Thus, this section presents the results of integration B, and C.

Effect of Experiments on Sophistication of Thinking

The quantitative summary of qualitative results (Step 3) presented in Table 4.25 below, displays the percentage of total responses in which each theme was found to occur, for a given question and experiment. For experiment 1, the percent count of the theme – concrete-intuitive physics was nearly twice that of theme – canonical physics, in learners’ responses, before the experiment (PreQ). Thus, students’ thinking was dominated by concrete-intuitive knowledge elements. However, after the experiment, the percent count of theme – canonical physics was found to increase by 71%, while the occurrence of concrete-intuitive KEs, decreased by 108%. The net gain (PostQ minus PreQ) for experiment 1, was therefore, +71% in terms of Canonical KEs, - 108% for concrete-intuitive KEs, and - 7% for abstract-intuitive KEs, which means that students’ sophistication of thinking increased overall, after the experiment. The gains for experiments 2, 3, and 4, were obtained by taking the average of gains observed for Post Q1 and PostQ2.

Table 4.25 Percent count of theme occurrences in total responses by question and experiment

Quan Summary % count	Experiment 1			Experiment 2			Experiment 3			Experiment 4		
	CP	I1P	I2P	CP	I1P	I2P	CP	I1P	I2P	CP	I1P	I2P
PreQ	58	113	12	96	13	39	76	8	11	67	13	18
PostQ1	129	5	5	64	64	5	63	21	24	43	32	8
PostQ2				132	15	40	63	13	13	52	52	0
Net gain	71	-108	-7	0	25	-18	-13	9	8	-18	30	-14
Qual Summary	Unclear notions about shock			Difficulty making canonical interpretations; less attention to area			Difficulty interpreting nonlinear graphs			Difficulty interpreting nonlinear graphs		

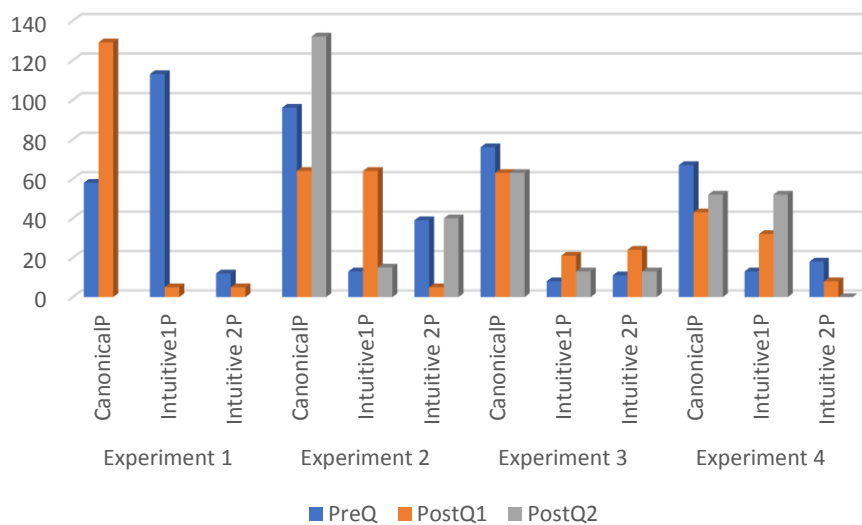


Figure 4.23. Percent distribution of themes of sophistication for PreQ, Post Q1, and PostQ2.

Figure 4.23 presents a visual summary of Table 4.24. Experiment 2 was found to yield no net gain in percent count of canonical KEs in learners' thinking, after the experiment, and Experiments 3 and 4 yielded net negative gains with respect to canonical knowledge. Concrete-intuitive KEs were found to show positive gains for Experiments 2, 3, and 4, while abstract-intuitive KEs showed negative gains for Experiments 2 and 4, and positive gains for Experiment 3. Thus, apart from Experiment 1, the experiments were not helpful in enhancing learners' canonical knowledge, but their abstract intuitive ideas seem to have been replaced, presumably, by concrete-intuitive ideas, at least for experiments 2 and 4. Qualitative analyses of learners' responses, also revealed 'hindrances' in learners' sophistication: which mainly was learners' difficulties in discriminating between non-linear and linear graphs. Specific connections and missed connections were revealed by network analyses, which provided a detailed map of learners' cognitive connections (co-occurrences, rather than occurrences). TA interactions could be one way of overcoming these hindrances, by collecting contextual data during inquiry activities.

Figure 4.24 presents a visual comparison of canonical connection strength (CCS), transitional connection strength (TCS), concrete-intuitive or intuitive 1 connection strength, abstract-intuitive or intuitive 2 connection strength, centrality or weight of canonical knowledge elements (K1C), centrality or weight of concrete-intuitive knowledge elements (K2C), and centrality or weight of abstract-intuitive knowledge elements (K3C), in learners' thinking, derived from the quantitative results of Step 6. Thus, the information made available by mixed methods, and quantitative results of network analyses, about individual learners' thinking, is more individualized, detailed, and rich, unlike qualitative results, which

gave approximate, and aggregated information, which is still indicative of effectiveness of activities, but not specific enough for initiating remedial measures.

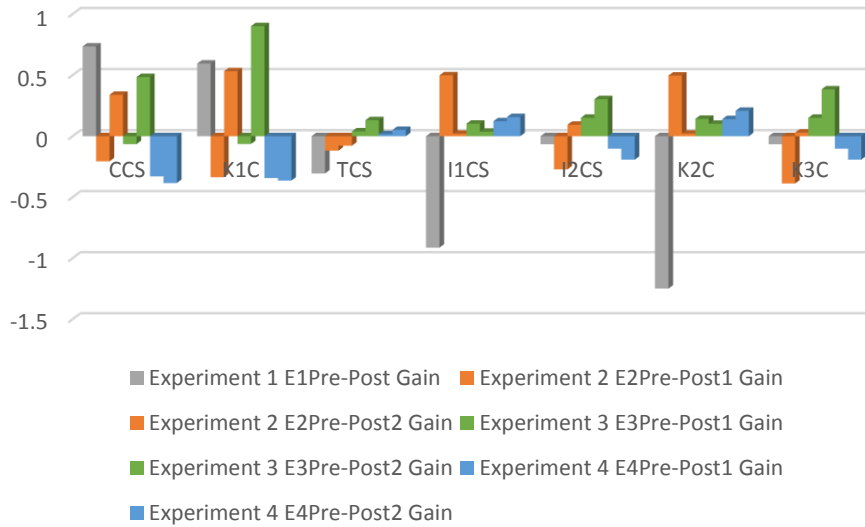


Figure 4.24. Gains in connection weights for Post1-PreQ, and Post2-PreQ, for Experiment 1(Charcoal grey), Experiment 2 (Orange), Experiment 3 (Green), and Experiment 4 (blue).

In figure 4.24, the two side-by-side columns of same color, for each category, represent the gains in sophistication, obtained by subtracting connection strengths for postQ1 and PreQ, and PostQ2, and PreQ. Only one column is found for Experiment 1 (grey color), because there was only one Post question. Positive gains in CCS, and K1C, and negative gains in other measures, is theoretically an ideal goal for sophistication of thinking. Figure 4.24 reveals that Experiment 1 yielded positive gains in CCS, K1C, and negative gains in TCS, I1CS, K2C, and, I2CS, and K3C. Thus, network analyses enhanced the findings of Figure 4.23, obtained from qualitative results, by showing explicitly, that Experiment 1 was effective in strengthening connections between canonical knowledge elements. Experiment 2 (orange color) yielded positive gains in CCS and K1C, for PostQ2, but negative gains for PostQ1. There were also positive gains in I1CS, and K2C for both questions (an indication of increase in unsophistication), and I2CS and K3C for PostQ2. However, I2CS and K3C showed decreased connections for PostQ1.

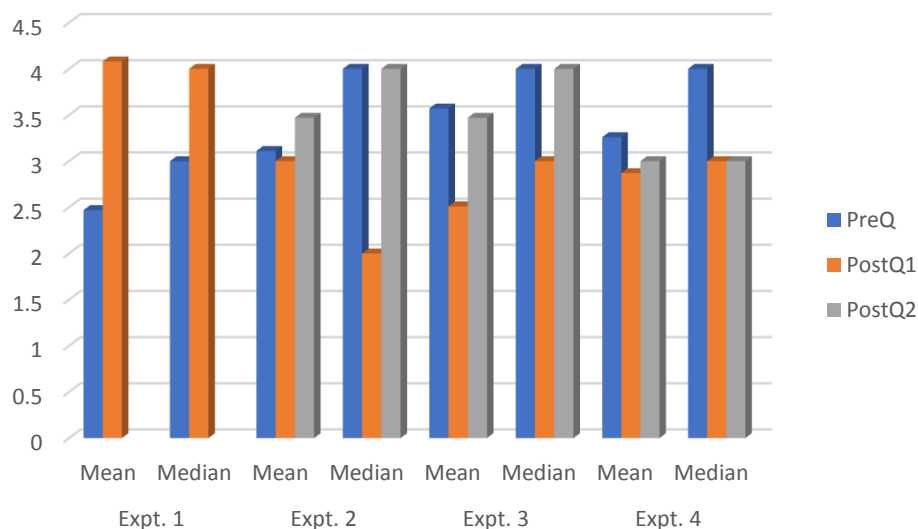


Figure 4.25. Comparative chart of mean and median sophistication levels for PreQ (blue), PostQ1 (orange), and PostQ2 (grey), for Experiment 1, Experiment 2, Experiment 3, and Experiment 4.

Figure 4.25 presents a comparison of mean and median sophistication levels, which summarize all the results presented in Figure 4.24, and will be used to compare the overall sophistication gains realized by each of the four experiments. The median sophistication levels remained unchanged for Experiments 2 and 3, for PostQ2, but a decrease in sophistication levels was found for PostQ1. For Experiment 2, this decrease could be because PostQ1, asked learners to describe how they changed the capacitance, and learners chose to refer to the concrete entities changed by them, without referring to the physical quantities changed by them, in turn. For Experiment 3, PostQ1 asked learners to classify light bulb as ohmic or nonohmic, and justify their answer based on the results of their experiment, and many learners made incorrect classification, because of their misinterpretations of linearity of graph. Thus, even if the decrease in PostQ1 S-levels is ignored, Experiment 2 did not yield any net sophistication in levels of thinking.

Experiment 3 can be said to have yielded a net decrease in sophistication levels, just like Experiment 4, which yielded a net decrease in S-levels, for both questions. An examination of Figure 4.24 however, reveals that there was one positive gain for Experiment 4 – decrease in abstract-intuitive connection strength, and centrality of abstract intuitive KEs in learners' thinking, after the experiment. All other measures of sophistication however, yielded negative gains in sophistication – decrease in CCs and K1C, and increase in TCS, I1CS, and K2C. Thus, Experiment 1, which did not involve any graphical outputs, was the only Experiment, which yielded tangible sophistication gains.

Validity and Reliability

I will draw upon the results of t-tests (Table 4.26), conducted to test for statistically significant differences between mean epistemic networks of Low, Medium, and High score groups on five quantitative assessments – EBAPS, A3 Home Work Quiz, Lab Quiz, Labs, and Test1, and the results of correlational analyses (Table 4.23), to provide evidence of validity of the ENA- based assessment explored in this study. The assessments from the course, used for correlation purposes, had proximal similarity with respect to time, except Test1, which took place after two weeks. The content covered was the same for ENA assessment, A3Quiz, Lab Quiz, and Labs. Items related to content, were also a part of Test1. Thus, Test 1 scores, were used, like other assessments, to check whether ENA assessment could discriminate between groups derived from other performance indicators used in existing assessments of the course. High correlation between ENA scores and other scores, was not expected, since the theoretical argument was, that ENA would be able to assess a construct – ‘connections in learners’ thinking’, which other assessments fall short of measuring, and that it would be able to do so, with a large enough sample which mimics the sample size presented by studio lab environments, unlike the small sample sizes, which typify qualitative assessments. Thus, achieving convergent validity was not a goal of validity testing, since there were no existing assessments, which measured the same construct – ‘cognitive connections’ directly.

Sophistication of thinking, however, is a broad enough construct, which includes conceptual understanding, the primary assessment objective of A3Quiz, Lab Quiz, and Test1. Hence, concurrent validity, by way of the ability of ENA assessment to discriminate between groups of low and high scorers on other assessments, was the goal of validity testing. Discriminant validity, was also a goal, by way of establishing, that there is only low correlation, if any, between ENA scores, and other scores, and therefore, they do indeed measure different constructs. Lab scores, although included in course grades, are not based on specific assessment rubric as such, and while the scores are expected to reflect learners’ understanding of their findings from the experiments, for practical considerations, the grading is flexible. Hence, correlations with Labs, did exist in some instances (experiments, and/questions), although theoretically, there should be none. Since EBAPS is an assessment which focuses on sophistication in learners’ thinking, but not conceptual understanding, and provides evidence of learners’ level of expertise with respect to their prior beliefs about physics content and learning, which are theorized to influence learners’ responses in-context, it was included in t-tests, and correlational analyses. However, all learners in this sample, demonstrated low levels of sophistication as assessed by EBAPS, and therefore, results of correlation with respect to EBAPS, may not be reliable.

Table 4.26 Summary of t-test Results for Significance of Differences in Epistemic Networks between Low, Medium, and High scorers on Five Quantitative Assessments

	Expt.1			Expt.2			Expt.3			Expt.4		
	PreQ	Post1	Post2	PreQ	Post1	Post2	PreQ	Post1	Post2	PreQ	Post1	Post2
EBAPS										M-H	L-M	
A3Quiz	L-H	L-M					L-M	L-M		L-M		L-M
Quiz4				L-H			M-H	L-H	M-H	L-H		
Labs							L-H	L-H			L-H	
Test1							L-M		M-H			

Table 4.26 presents a summary of t-test results which were obtained through Step 5 (Table 4.22). ENA assessment could discriminate consistently between low-medium and/ low-high score groups for all four experiments, except Experiment 2; between low-high score groups of Quiz4 (lab Quiz) for experiments 2, 3, and 4; between low-high groups of Labs for Experiments 2, 3, and 4. Differences in epistemic networks between groups were statistically significant, for at least three assessment instruments, including EBAPS and Test 1 scores, for Experiments 3 and 4, which could also mean, that the questions (items) of ENA assessments for these experiments had higher discriminant indices than experiments 1 and 2. It could also be, that the difficulty indices of items of Experiments 3 and 4, were higher, than Experiments 1 and 2, although all questions were derived based on the content of readings, lab sessions, and inquiry tasks, the difficulty levels of content covered were different.

Table 4.27 presents a summary of results of correlational analyses (Spearman' rho correlation coefficient) between sophistication scores of ENA assessment, and five quantitative assessments used for validation purposes. Specific values can be found in previous section. The Table shows pairs of assessments which showed weak to medium correlations (coefficients between 0.3 and 0.4). Positive correlations could be expected between K1C, and CCS, and negative correlations between K2C, I1CS, K3C, and I2CS. Correlations with TCS, could be either way, and therefore, not reliable indicators of validity. Positive correlations of significance, were found between K1C and Test1 (PostQ, Experiment 1), and between CCS and A3Quiz (PostQ2, Experiment2). Negative correlations of significance were found between K2C, I1CS, and Lab4, Lab Quiz (PostQ1, Experiment 3). Thus, there was not much evidence of correlations between ENA measures and other assessments, which could be construed as confirming that these two sets of measures measured dissimilar constructs.

Table 4.27 Summary of Correlations between Sophistication Measures and Quantitative Assessments

	Positive Correlations	Question	Negative Correlation	Question
Experiment 1	Test 1 – K1C	PostQ	EBAPS – K2C EBAPS – I1CS	PostQ PostQ
Experiment 2	Test 1 - TCS A3Quiz - CCS Lab 4 – K2C Lab4 – I1CS	PostQ1 PostQ2 PostQ2 PostQ2	Lab 4 - TCS	PreQ
Experiment 3	EBAPS – I1CS	PostQ2	A3Quiz - TCS Lab Quiz – K2C Lab Quiz – I1CS Lab 4 – K2C Lab 4 – I1CS	PostQ1 PostQ1 PostQ1 PostQ1 PostQ1
Experiment 4	Lab 4 – I2CS	PostQ1		

Discriminant validity presented by Table 4.27, and concurrent validity, presented by Table 4.26, are indicative, rather than conclusive, because of lack of consistency of findings. Thus, more studies, with other inquiries, and development of extensive code libraries which reflect further differentiation of code categories by knowledge representations, are required for achieving construct validity.

Reliability. Reliability of the assessment is primarily resting on its repeated administration to four different Experiments, with consistency. Inter-item correlation and Cronbach’s alpha were also calculated, to test for internal consistency. The pairs of related measures were – (i) CCS, and K1C, both of which estimate the strength and weight of canonical connections, and canonical knowledge elements respectively; (ii) I1CS and K2C, both of which estimate the strength and weight of concrete-intuitive connections, and concrete-intuitive knowledge elements respectively; and (iii) I2CS and K3C, both of which estimate the strength and weight of abstract-intuitive connections, and abstract-intuitive knowledge elements respectively. For computation of correlation and Cronbach’s alpha, TCS, strength of transitional connections was included in the first category, along with CCS and K1C, because it shares the occurrence of canonical knowledge elements, with the other two measures.

Results of these analyses, are presented in Tables 14.1 to 14.4 of Appendix 14; one Table per Experiment. Inter-item correlation, and alpha values were consistently high for the three groups of pairs specified above. As expected, negative correlation of CCS and K1C, was however found with TCS, for Experiments 1, 2, and 3. Responses to some questions did not have TCS connections, and therefore, TCS was excluded by SPSS, from the calculation of Cronbach alpha. Thus, the assessment measures were found to have internal consistency reliability. TCS, by its composition, is not a reliable measure of sophistication, but indicator of the transitional status of learners’ thinking, since K1, K2, K3, co-occur in

these connections, and it can be looked upon as learner's state of transition from unsophisticated to sophisticated thinking.

Inter-rater agreement. Inter-rater agreement on the qualitative coding for all Experiments and questions, was 96% on an average. (Table 15.1, Appendix15)

Summary of Mixed Methods Findings

In this section, the findings of previous sections were combined and compared, to answer research question 4. Experiments were compared for their effect on sophistication of learners' thinking, by drawing upon the results of qualitative analyses, network analyses, and quantitative analyses. It was found, that Experiment 1 had the most positive gains in terms of increasing learners' sophistication levels, and connection strength of canonical connections. Experiment 2 yielded no net gain, and Experiments 3, and 4, yielded negative gains. The primary barrier, seemed to be learners' prior inexperience with interpretation of graphical outputs generated by real life instruments, and could be overcome by timely intervention by the TAs. The validity of the assessment, was found to be indicative, rather than conclusive, and further research is required for validation purposes. However, reliability of the measures proposed, was found to be good, and the assessment can be said to provide a reliable tool for the assessment of sophistication of thinking, and its validity can be enhanced through refinements of the coding process, and fine tuning of categories.

CHAPTER 5 CONCLUSIONS AND RECOMMENDATIONS

Research-based active learning environments (ALEs), commonly employed in introductory university physics courses, are designed to enhance the sophistication of learners' thinking – to enable the development of strong cognitive connections between relevant knowledge elements, their representations, and reality, such that they can be activated consistently, and coherently, when learners are presented with discipline-specific problems/tasks (Meltzer, and Thornton, 2012). Formative assessments are developed and administered before, during, and/ after instructional activities, to monitor and track the sophistication of learners' thinking (McDermott, 2001). Clickers – technology-enhanced classroom response systems, have gained popularity with instructors of introductory physics courses, to monitor learners' conceptual understanding. However, most existing assessments, typically target the presence or absence of knowledge, skills, epistemologies, beliefs, identities, values, and/expectations, or the KSEs, and therefore, fall short of capturing their co-occurrences, or the 'connections' made, missed, or the misconceptions between the KSEs found to occur in learners' thinking (NRC, 2001; Larkin, 2011). Researchers have used qualitative methods to examine learners' thinking and reasoning process, to identify learners' epistemic frames (Tuminaro, 2004, Bing, 2008), and to model the connections in their thinking, in the context of virtual learning environments (Arastoopour, et. Al, 2012). To implement these techniques in studio physics contexts, there was a need to identify appropriate technologies to collect, analyze, model, and quantize qualitative data on learners' thinking, to execute an assessment process.

The purpose of this study was to explore, how technology can be leveraged, to enable the assessment of sophistication of learners' cognitive connections, during inquiry experiments, conducted collaboratively, in studio physics setting, of an introductory university physics course. The assessment framework can be extended to high school and more advanced undergraduate courses taken by physics and engineering majors. Such assessments will provide tangible information about the impact of instructional activities on sophistication of cognitive connections, and therefore, will be more effective in the evaluation and refinement of instruction. The proposed assessment mechanism will also enable instructors to get a deeper and broader view of learner differences in framing, reasoning, and connecting, and thereby to monitor the inclusivity of their instruction. In this study, I explored the interfacing of four commonly used and/open source Technologies to develop the assessment. Social Media Technology, Twitter, was used to collect qualitative data; MS Office spreadsheet package, was used to qualitatively code the data; an open source package, ENA Web Tool 4.0, was used to model the coded data as networks and matrices; and an open source statistical computing software package, RStudio, was used to enable quantitative assessment of the matrix data.

Primary Finding

The study found, that the proposed assessment, which utilized the interpretive lens of Epistemic Network Analysis (ENA), was indeed suitable for formative classroom assessment in inquiry settings, and that such an ENA-based assessment can be realized, by implementing the following four basic steps, which form a framework of assessment, which I choose to call, QOT2SM (schematic is presented in Figure 5.1).

- (i) Collect Qualitative Observations of Thinking (QOT).
- (ii) Identify and Code the Elements of Sophistication and their occurrence counts (SE).
- (iii) Model the co-occurrences of SEs as Sophistication Network graphs and Matrices, (SN-G, SN-M).
- (iv) Compute Measures of Sophistication (SM).

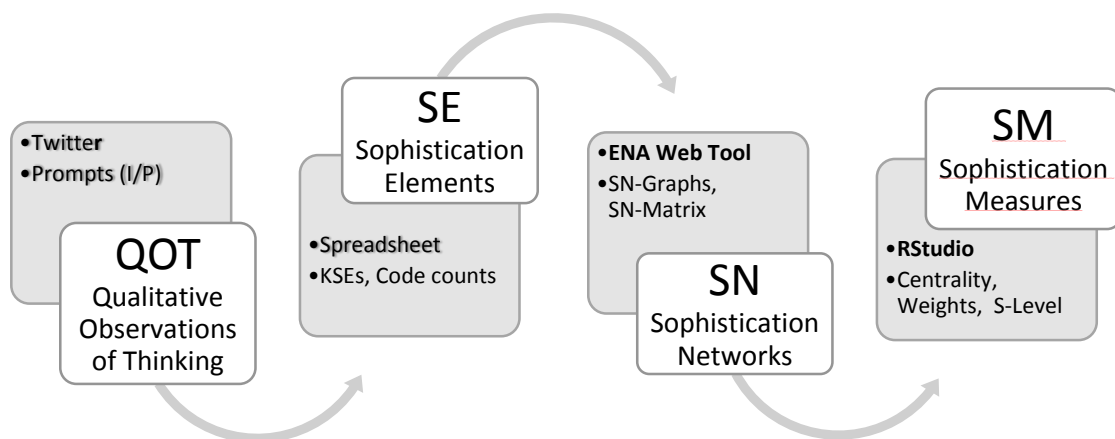


Figure 5.1. QOT2SM Framework of an ENA-based assessment.

The four blocks or components which executed each of the four steps listed above, are labelled by the assessment component, generated by the block, as follows:

- The QOT block of the framework generates qualitative observations of learners' thinking (QOT) as output, with 'Question prompts' serving as input, and 'Twitter' as the technology used to deliver the input and retrieve the output.
- The SE block generates sophistication elements (SEs) which characterized learners' thinking, as output, with QOT as the input, and Spreadsheet as the technology used to process the input – qualitative coding.

- The SN block generates sophistication networks (SNs), of learners' thinking, as SN-G sophistication network graphs, and SN-M sophistication network matrices, with SEs and their occurrence counts as input, and ENA Web Tool as the technology used, to process the input – network analyses.
- The SM block generates sophistication measures (SM) – centrality of sophistication elements, connection weights, and level of sophistication (S-level), of learners' thinking, as output, with SN-M as the input, and RStudio as the technology used to process SN-M – quantitative analyses and computations.

The study adopted a single-study, multi-phase, QUAL-QUAN (equal priority) mixed methods approach to design, by integrating qualitative and quantitative data collection and analyses at multiple points in the research study, to answer four research questions. Methods were mixed in the study, in such a way, that each analysis informed and/ helped develop the subsequent analyses conducted in this study. In this chapter, I will first present a summary and discussion of main findings by research question (RQ), and draw upon relevant literature, to highlight similarities and differences. Next, I will discuss implications of the study and its findings, for practice and research, in separate sections. The chapter will conclude with my closing thoughts, and an overview of the next steps I wish to pursue, to refine and expand, the exploratory findings of this study.

Discussion of Main Findings

As discussed in Chapter 2, the process up to step (iii) had been tested previously, in other contexts in virtual and classroom spaces, in other domains; but it was not previously implemented in inquiry contexts in physics (Arastoopour, et.al., 2016; Bodin, 2012). Out of the two main approaches used by researchers, to examine learners' epistemic frames – Epistemic Framing Analysis, and Epistemic Network Analysis, the approach of Epistemic Framing Analysis, is purely qualitative in nature, and utilizes traditional data collection instruments such as interviews (Lisin, and Elby, 2005), video recording (Lising, and Elby, 2005; Scherr, and Hammer, 2009; Tuminaro, 2004), and/non-digital artefacts of student work (Lising and Elby, 2005; May, and Etkina, 2002), which not only limit the sample size, but generate data which is cumbersome to process. So, while these studies, which utilized epistemic framing analyses, have all yielded useful information about learners' epistemic framing of inquiry, and/problem solving tasks in undergraduate physics, the data collection and analytical techniques employed by them for conducting epistemic framing analyses, are unsuitable for use as regular formative assessments, due to their inadequacy to handle large sample size (40 to 45) presented by studio lab environments. The second analytical approach used to examine learners' epistemic frames – Epistemic Network Analysis (ENA), has been used previously in virtual learning environments such as virtual internships in Engineering

(Arastoopour, et.al., 2016; Svarovsky, 2011), by drawing upon online chat discourse data, digital logs, and other digital artefacts of student work, to identify epistemic frame elements, e.g. knowledge, skills, values, epistemologies, and identities.

However, studio lab contexts in undergraduate physics often do not have readily available sources of digital data on student thinking. Also, the pace and structure of collaborative inquiry activities in these environments, make it difficult to design and administer formative assessments which can extract qualitative observations of learners' thinking in-context. Multiple-choice questions administered using Clickers are commonly used, to gauge student understanding of concepts explored through activities, before and/after activities. But the aggregated graphical output generated by Clickers, presents only a snapshot of percentage of students who chose a specific 'response' for each question, which tells us how many students thought or believed a 'correct' or 'incorrect' response to be a valid answer, but does not tell us what the students were thinking when they decided to choose a specific response. To uncover the epistemological underpinnings of learners' thinking, it is necessary to collect qualitative observations of thinking, and conduct Epistemic Network Analysis, as was done by Bodin (2012), to examine shifts in learners' thinking after completing a computational physics task. However, traditional interviews were used to collect data, thus limiting the sample size to six students.

The exploration of Twitter, in this study, to enable the collection of qualitative observations of student thinking, thus has contributed valuable knowledge towards the implementation of ENA for assessing sophistication of learners' thinking in inquiry contexts in physics, while also going a step further, towards quantization of sophistication, a necessary step, in enhancing the performance utility of an ENA-based assessment, for formative assessment in undergraduate physics. The aggregation and quantization realized in this study was that of learners' own articulations about the concept, and not only mirrored learners' canonical as well as intuitive understandings, but also further differentiated the intuitive understandings, to identify the 'source' or 'epistemological barrier' in learners' thinking. To realize this assessment of intuitive epistemological barriers, this study deviated from the traditional categorization of epistemic frame elements as knowledge, skills, epistemologies, values, and identities (Arastoopour, et. Al., 2016; Svarovsky, 2011; Bodin, 2012), by focusing on the knowledge and epistemological elements, and stratifying the knowledge category, to create canonical and intuitive knowledge categories, and further elaborating the intuitive knowledge category, as concrete-intuitive, and abstract-intuitive knowledge.

This conscious deviation from and adaptation of traditional ENA was driven by the findings of 'themes' which emerged from the data. Emergent themes seemed to resonate with Vygotsky's (1962) classification of concepts as scientific (canonical) and spontaneous (intuitive), and his theory, that both

concepts develop simultaneously, but in opposite directions, in learners' cognition. Scientific concepts develop from abstraction towards meaning, while spontaneous concepts develop from experience and meaning towards abstraction. Conceptual understanding happens when the formation of both concepts intersects and/ overlap (Vygotsky, 1962). Such adaptations are the primary benefit of using the ENA approach and have been made previously in the context of analyzing surgeons' handling of errors during surgery (Ruis, et. Al., 2017). Specific findings by question are presented and discussed below.

Findings: RQ 1

- *Learners' Epistemic constructions (E) of concepts inquired through four experiments covering Electric circuits, in studio lab settings in undergraduate physics, were found to differentiate in sophistication (SE), by the type of knowledge – canonical (K1), or intuitive, and then, by the source of knowledge - concrete-intuitive (K2), or abstract-intuitive (K3).*
- *Learners' epistemic constructions demonstrated predominantly canonical knowledge representations, before the experiments, but the simultaneous occurrence of intuitive knowledge elements, was found to either not change, or found to increase by differing extents, and numbers, after the experiments*

The goal of RQ1, was to find the dominant themes in learners' thinking, as they conducted four Experiments, inquiring concepts from 'electric circuits', in groups of three, during a 3-hour long studio lab session, in a non-Majors course on Electricity, Magnetism, Optics, and Modern Physics. It was achieved by qualitatively coding Twitter responses (direct messages) of learners, to pre-and post-questions sent to them, before, and after each of the four Experiments. Findings revealed that learners' thinking in the given context, reflected four broad themes: (i) Epistemic construct (E) - the concept learners believed or interpreted as being referred to in the question prompt; (ii) Canonical physics (K1) - canonically appropriate explanations, descriptions, justifications; (iii) Concrete-intuitive physics (K2) - explanations, descriptions, justifications which used concrete entities (examples) from experience, or from course readings, to articulate the responses to question prompts, which demonstrated vague or indirect references to canonically appropriate constructions; and (iv) Abstract-intuitive physics (K3) – explanations, descriptions, justifications which were irrelevant, ambiguous, inappropriate or vague, and seemed to originate from prior misconstructions of related canonical ideas, and/relationships.

This finding is broadly consistent with findings of previous studies, that novice learners such as those taking introductory physics courses, demonstrate context-sensitivity of knowledge (diSessa, 1998; 2016; Hammer, and Elby, 2003; May, and Etkina, 2002). In this study, learners' responses were found to demonstrate canonical knowledge before experiment, and intuitive knowledge after the experiment. Post Q2s were designed to ask learners to draw inferences from inquiry experiments, about relationships, which they were found to have articulated canonically, before the experiment. But their articulations after

the experiments were found to refer to concrete experiences, without abstracting them to the underlying concepts. While this finding about the nature of learners' responses, agrees with previous works, the knowledge categories identified in this study, differ from previous works utilizing the lens of epistemic network analysis, which typically sort codes into categories of knowledge, skills, epistemologies, values, and/identities (Bodin, 2012; Arastoopour, et.al., 2012; 2016).

The broad categories of knowledge, and epistemologies, were retained, like previous works in epistemic network analysis, but the knowledge category was further differentiated, first, as canonical and intuitive knowledge, following diSessa's (2016) broad categorization. Then, the intuitive category was further elaborated, based on the source of knowledge, adapting diSessa (1988; 2016), and Vygotsky's (1962) approaches to classification of concepts, as discussed in Chapter3. I contend, that while the classification adopted in this study, may not neatly conform to theory, it attempts to bridge the gap between research and practice, by providing a categorization, which is not distant from the instructor's routine assessments, and yet, draws upon theories of concept formation. The rationale for this adaptation, is that ENA was designed to afford openness to interpretation of SEs (Shaffer, 2017), and it is this feature, which makes it a good fit for developing customized formative assessments. Researchers, have previously utilized this feature in analyzing gaze coordination (Andrist, et.al., 2015); and assessing intraoperative performance of surgeons, by modeling the way they coordinated psychomotor, procedural, and cognitive skills, to manage errors during surgery (Ruis, et.al., 2017). The elements of sophistication in these works were defined in relation to the context, as was done in this study.

Findings: RQ 2

- *The structure of cognitive connections evidenced in learners' responses to probing questions, asking them to describe-explain-justify concepts and relationships inquired through four experiments covering Electric circuits, in studio lab settings in undergraduate physics, was found to differentiate in sophistication, by the Elements of Sophistication making up the connection –CC. or canonical connections (E-K1, K1-K1), TC, or transitional connections (K1-K2, K1-K3), I1C, or intuitive1connections (E-K2,K2-K2), and I2C, or intuitive 2connections (E-K3,K3-K3,K2-K3).*
- *The structure of connections, were found to differ significantly, before and after all four experiments, suggesting that learners' epistemic constructions of concepts, shifted significantly, after the experiments, and that the experiments were likely to have been instrumental, in effecting the shifts.*
- *The structure of connections differed significantly by Section of the course, 50% of the time, i.e., for experiments 2, and 3, (2 out of 4 experiments). Thus, TA effect can also be considered as a factor which contributed to shifts in thinking.*

- *The structure of connections differed significantly, by students grouped on scores of Lab Quiz, Labs, and A3HWQuiz, for Experiments 3, and 4. Thus, concurrent validity could not be established conclusively, between ENA networks, and A3HW Quiz, Lab Quiz, and Labs, since there was no consistency of findings. However, this finding lent support to the rationale of inadequacy of existing assessments, to measure sophistication of connections in thinking.*

RQ2 sought to identify the mean patterns of connections in learners' thinking, and differences in connections, of student groups, sorted on the scores of existing assessments, administered in the same context. It was addressed by conducting network analyses of learners' epistemic networks of sophistication, (SN), modeled from the co-occurrences of sophistication elements (SE), as revealed by qualitative coding of learners' Twitter responses (QOT) to questions for each experiment. Findings revealed that learners' thinking in the given context, reflected four broad categories of connections: (i) Canonical (CC) – between epistemic construct and canonical knowledge elements; (ii) Transitional (TC) – between canonical and intuitive knowledge elements, which typify a transition from one stage of concept formation to the other; (iii) Intuitive 1 (I1C) – between epistemic construct and concrete-intuitive knowledge elements, which typify a more sophisticated level of intuitive connections as compared to the next category; and (iv) Intuitive 2 (I2C) – between epistemic construct and abstract-intuitive knowledge elements, and between intuitive knowledge elements. The categorization of connections by knowledge type, and/source, has not been done previously, and was operationalized in this study, by the adaptation implemented, in the categorization of knowledge elements described in the previous step. This operationalization was required for further quantization of the data, and to generate scores of sophistication.

Svarovsky (2011) described learning gains in the context of a virtual engineering internship, as significant increase in references (connections to) to engineering knowledge, skills, epistemologies, identities, and values, in pre-post, and extended time lapse (3 months) interviews, asking learners to explain concepts, give opinions, and/ justify design choices. Bodin (2012) described the findings of a study examining shifts in learners' epistemic frames during a computational physics task, as change in epistemic framing from a computer modeling task, with expectancies about learning programming, before the activity, to a physics task, during the activity. This study took a different approach to characterize epistemic frame connections, which involved differentiation of 'knowledge' category. This methodological decision was based on two factors – (i) the type of questions that could be asked in the context examined, predetermined the nature of responses expected viz. different types of knowledge (ii) the operationalization of quantitative sophistication measures envisaged in this study (next step),

necessitated specification of connections by knowledge type in this step, which in turn, necessitated elaboration of knowledge category, in the previous step.

However, the methodology adopted for qualitative and quantitative (t-tests) characterization of these connections, through examination of network graphs, was consistent with previous works utilizing the lens of epistemic network analysis (Bodin, 2012; Arastoopour, et.al., 2012; 2016). Qualitative, and quantitative analyses of network graphs, addressed two parts of RQ2. The first part (first three bullet points above) was answered descriptively, and statistically (T-test for significance of differences in mean) by examining pre-post mean and individual network graphs corresponding to responses for each question and experiment. The finding, that structure of connections was found to shift, or change, after the experiment, was consistent with other works (Bodin, 2012; Svarovsky, 2011; Arastoopour, et.al., 2012; 2016). Apart from this broad finding, there were specific findings revealed from analyses of each experiment, and question, as discussed elaborately in Chapter 4. Two major specific findings of this study were – (a) Significant differences in mean epistemic networks by Section of Course, for 2 experiments; and (b) Evidence of ‘epistemological barriers’ with respect to interpretation of linearity and non-linearity of graphs, for the same two experiments.

These findings show that even in the context of ALEs such as studio physics, whose effectiveness has been validated through multiple studies (Beichner, 2008), the instructor effect or in this case, the TA Effect can be a significant factor in differentiating learners’ epistemic frames. Especially so, in scenarios, or contexts such as experiments 3 and 4, which, were rich in graphing tasks, but learners’ ‘epistemological barriers’ to interpreting those graphs correctly, remained hidden, and therefore, unaddressed; thereby significantly affecting the nature of learners’ responses. The pre-responses were canonical, but the post-responses were intuitive, confused, and hazy, because although they were found to be situated in learners’ inquiry experience, they were hindered by learners’ difficulty of connecting to (transferring) their knowledge of graphing from math, in inquiry contexts in physics. While both findings ‘a’ and ‘b’ above, have not been explicitly examined previously using ENA, they are consistent with findings of other works in PER, and Science Education, as explained below.

Frequency of TA-student interactions, not the length of interactions, has been found to have positive effect on outcomes of student engagement in physics labs (Stang, and Roll, 2014). Having a trained, and supported TA was found to boost scores of above-average students significantly, and was also found to be associated with students’ persistence into the next course of two-course introductory chemistry sequence, for science and engineering majors (Philipp, Tretter, and Rich, 2016). Student strategies of graph interpretation have been found to be largely context dependent and domain specific, in an investigation of 385 first-year university students, when asked to explain their interpretations of graphs

in three different domains, including physics. In physics, the difficulty seemed to originate in students' expectation that physics involved the use of formulas (Ivanjek, et.al., 2016). Two categories of difficulties were identified in an extended study examining common errors in graph interpretation, of several hundred university students taking a lab-based introductory physics course – (i) difficulty in connecting graphs to physical concepts, and (ii) difficulty in connecting graphs to the real world (McDermott, Rosenquist, and Van Zee, 1987). Thus, the QOT2SM assessment implemented in this study, could detect these two commonly found (a and b above) and extensively researched, learning and teaching problems, in physics, and this could be viewed as a validation of the instrument.

The second part, of RQ2, was addressed by testing for statistical significance of differences in mean networks of learners, grouped by their scores on four existing assessments of the course, and a standardized survey instrument, assessing learners' epistemological beliefs about physical sciences. The findings (fourth bullet point above) were expected to provide concurrent validity, if statistically significant differences were found, since these measurements (tests) occurred close together in time (were conducted concurrently) and covered the same content. However, there was no consistency in differentiation of epistemic networks by existing assessments, across experiments and questions. Out of 9 comparisons conducted for each existing assessment, by experiment (3 groups of low, medium, high, by three questions), only networks corresponding to Experiments 3, and 4, showed significant differences by A3 HW Quiz, and Lab Quiz, for 33% (6 times out of 18) and 39% (7 times out of 18) of the total comparisons, respectively. Across experiments, Experiment 3 showed significant differences, for 39% (14 out of 36), of total comparisons. Thus, Experiment 3, and Lab Quiz, could be said to provide concurrent validity to ENA assessed as network graphs, at least 39% of the time. It should be noted, however, that the lack of concurrent validity, could also be viewed as corroborating the discriminant validity, established from the findings of RQ3, discussed in the next sub-section. As discussed in Chapter 4, it could also, point towards issues with existing assessments themselves, in terms of their discrimination indices being low, or, it could be an affirmation of the rationale, that existing assessments fall short of truly assessing connections in thinking, as are explicitly assessed by network graphs and matrices.

Findings: RQ 3

- *The level of sophistication of thinking evidenced in learners' responses to probing questions, asking them to describe-explain-justify concepts and relationships inquired through four experiments covering Electric circuits, in studio lab settings in undergraduate physics, was found to differentiate by the strength of connections making up their epistemic networks – Level 1, with strong intuitive 2 connections, Level 2, with strong intuitive 1 connections, Level 3, with strong transitional*

connections, Level 4 with strong canonical connections, and Level 5, with very strong canonical connections.

- *Weak or no correlations were found between the measures of sophistication developed to assess sophistication of thinking using the ENA-base assessment, and the scores of EBAPS, A3 HW Quiz, Lab Quiz, Labs, and Test1, thereby establishing Discriminant validity of the proposed assessment.*

The objective of RQ3 was to determine the levels of sophistication in learners' thinking (SM), in the same context, and it was addressed by conducting quantitative analyses on the network matrix datasets (SN-M) derived through network analyses of learners' responses to each question prompt, for each of the four experiments. Thus, 8 sophistication measures (SM), were computed as follows: (i) 3 measures of node centrality – K1 Centrality, K2 Centrality, K3 Centrality, which measured the centrality of each of these knowledge elements in learners' thinking; (ii) 4 measures of connection weight – CCS, or canonical connection strength, TCS, or transitional connection strength, I1CS, or intuitive1 connection strength, I2CS, or intuitive 2 connection strength, which measured the weight or sum of strengths of each of these connections; and (iii) S-level, or level of sophistication, which was determined by using criteria to compare the relative strengths of each of the 4 types of connections, in learners' thinking. Learners were found to demonstrate five levels of sophistication of thinking, which could be characterized by the strength and type of connections found in learners' thinking. This operationalization was done on an exploratory basis, and the goal was to examine their validity and reliability. Because of its developmental nature, there is no precedence existing in literature, with respect to studies describing the types of connections found in learners' thinking, quantitatively.

The second part of RQ3, was addressed by conducting correlational analyses using SPSS, between the 3 scores of node centrality, and 4 scores of connection weight, and learners' scores on four existing course assessments, and one standardized beliefs survey, which found that there were no significant correlations, as measured by Spearman's rho-coefficient, whose values were equal to or less than 0.3. This low correlation with existing assessments, essentially means that the ENA-based assessment, and existing assessments, measure different attributes, thereby establishing its discriminant validity.

Findings: RQ 4

- *The study found positive sophistication gains for Experiment 1, no net gains for Experiment 2, and net negative sophistication gains, for Experiments 3, and 4.*
- *The study found that the ENA-based assessment was reliable, but its validity needs to be tested through further studies, and refinements of outputs derived from each block of the assessment. Only discriminant validity could be established, for want of comparable assessment instruments.*

RQ4 sought to compare the effects of four experiments which formed the context of this study, on the sophistication of learners' thinking, and, to analyze the validity and reliability of the measures of sophistication, developed through this study. It was addressed using mixed methods, to integrate the findings of qualitative, network, and quantitative analyses. Each block of QOT2SM framework, transformed or processed the qualitative observations of thinking (QOT), to yield substantive information about sophistication of thinking, but each output differed significantly from the other. The second block transformed QOT by qualitative coding, to reveal dominant elements which characterized sophistication of thinking (QOT2SE); but it did not account for co-occurrences of more and less sophisticated elements. It is important to capture this information, for revealing which co-occurrences have more weight, and which elements are more central in learners' thinking. The third block, transformed QOT to network graphs (QOT2SN-G), and thereby, enhanced the information provided by the previous block, by (a) counting the co-occurrences, and (b) presenting a versatile visual representation. Most applications of ENA in education, have utilized the qualitative quantization feature of the third block, to conduct graphical network analyses, and t-tests (Arastoopour, et.al., 2016; Bodin, 2012).

Through this study, I sought to examine, whether, by going one step further, ENA can be utilized effectively, as a regular formative assessment instrument, in inquiry settings, in undergraduate physics. The fourth block transformed QOT to sophistication measures (QOT2SM), and provided numeric information, about the network graphs generated by the previous block, thereby enhancing the utility of this information for regular formative assessment in classroom settings. The graphical representations, did yield rich information as well, which overcame the limitation of purely qualitative techniques of the previous block. However, the network graph analyses still had an element of cumbersome qualitative analyses, which needed to be quantized, to enhance their performance utility, as regular formative assessments in inquiry contexts in physics. The quantization realized through this study, has achieved this goal. One study (Rus, et.al., 2016) which used data from previous ENA studies on two virtual engineering internships, has taken a similar exploratory approach, as taken by this study, to find ways of automating the qualitative coding process. Rus et. al. (2016) acknowledge the need for quantization, and support the argument made by this study, that quantization and aggregation are essential features of any automated classroom assessment system.

Sophistication Gains. As discussed elaborately in Chapter 4, only one experiment yielded net positive gains in sophistication. The zero or net negative gains found for other three experiments, brought out epistemological issues, which were different for each experiment, but one common theme did emerge. Students had difficulty in connecting to 'appropriate math' from their repertoire of knowledge, about linear and non-linear graphs, in real world contexts of interpreting graphical outputs generated by

laboratory equipment, while conducting inquiry experiments to investigate ‘electric circuits’. These difficulties have previously been found

Reliability. This study has shown that the computation of SMs can be realized, by adding one more block, and that these measures were reliable, not only because Cronbach alpha values were high, but because the assessment was implemented repeatedly, and successfully, across four different experiments, whose content ranged in difficulty level from easy (body resistance), and average (capacitance, and ohmic-nonohmic materials) to complex (RC circuits).

Validity. Although the validity of the ENA-based assessment as an instrument, could not be established conclusively in this study, it should be noted that ENA has been used extensively to model and assess coordination of performance and/behavior, during complex problem-solving activities not only in the field of education, but investigative journalism, primary care, urban planning, and trauma care. Thus, the modeling and assessment capabilities of ENA have been validated through wide ranging applications in diverse fields. However, the specific coding schemes, and interpretive mechanisms chosen, need refinements through further studies, to establish validity of their applicability for assessments in specific contexts and fields, as in inquiry contexts in undergraduate physics.

The ENA-based framework developed and tested in this study, offers a structure, which can be used to develop customized assessments tailored to suit assessment needs of individual instructor, based on instructor’s expectations of sophistication, learning goals, the nature of content, and the nature of instructional activities, (a) by substituting technologies, (b) by using alternative measures to obtain outputs of one or more blocks, and (c) by using only qualitative, only quantitative or mixed methods measures. The possibilities of adaptations will be discussed in the next sections, under implications for practice, and research.

Implications for Practice

I will discuss here, what the findings reveal about the lab activities and the assessments used in the course, and make specific recommendations for modulating these two components, such that their effectiveness in sophistication of learners’ thinking, and assessment of thinking, can be enhanced. These recommendations are based on two research-based, widely accepted principles which frame the design of effective active learning environments in undergraduate physics. First, is that formative assessments must be an integral component of active learning environments; Second, effective formative assessments make student thinking visible, and their purpose should be to give feedback to instructors and to learners on their learning, to adjust or modify instruction, and to adjust ways of learning, respectively (Kober, 2015).

I make four specific recommendations for enhancing the pedagogic, instructional and assessment value of existing formative assessments. First, is for modifications to the administration of Lab Reports; second, is for modifying and aligning the Lab Quizzes with contents covered by experiments; third is for

enhancing the formative aspect of Homework Quizzes, and fourth is, for incorporating Twitter as a regular feature of the course, to enable peer to peer interactions, and student-TA interactions, which can have tremendous usefulness for social construction of canonical knowledge. These recommendations are discussed below, not in a specific order, but by their relevance to the discussion of issues with existing activities, and assessments, and suggested solutions.

Studio Physics (Labs)

PHYS222 course which formed the setting of this study, utilizes the research-based studio physics approach to instruction, which typically combines the lecture, lab and recitation components of a traditional physics course. Thus, the instruction delivery format is – minimal lecture, cooperative learning through labs or inquiries, and focus on student reasoning and presentation. While the lecture component is minimized, the feedback component is maximized through student-TA interactions, facilitated by the student seating design adopted in Studio physics classrooms (Figure 5.2). Formative assessments which make student reasoning visible, complemented by TA/instructor feedback, is thus, a defining feature of studio physics classrooms (Beichner, 2012). In a recent study which observed and coded thirteen instructors teaching one specific implementation of studio physics – the SCALE-UP (student-centered active learning environments in undergraduate physics) model, at two universities, on the percentage of class time spent on four activities, ‘lecture’, ‘feedback’, ‘clickers’, and ‘small group work’, it was found, that time spent on group work ranged from 60 to 70%, time spent on lecture, was 21% on an average, on clickers, from 11 to 21%, and on feedback, from 25 to 30%, on an average (Wilcox, et.al., 2016).

Timing issues. From my observations of the studio sessions, in this study, the studio labs in PHYS222, seemed to follow the same overall model, with a major difference – substitution of Clicker assessments by Lab Quizzes, in paper format, delivered before the group work began. While Clicker questions are generally presented at intervals, which coincide with the beginning of each group activity, and the responses are immediately analyzed and discussed, the lab quizzes are administered only once, and are graded after the studio session. This substitution makes the lab quiz more summative in nature, than formative. Also, lab reports, which embody the other formative assessment instrument, are submitted much later, and the feedback given by TAs is subsumed by their evaluative grades (Kober, 2015; NRC 2001). When feedback and evaluation are given simultaneously, students tend to process the grade, and ignore the feedback. Much of the feedback given by TAs on the lab reports, was useful in triangulating my interpretations of learners’ responses to PostQ, and thus contained useful qualitative evaluation of student work. If the same feedback were given in-context, its value for enhancing learning, would improve significantly. It would also motivate students to interact with TAs, in response to their feedback, since students would be interested in enhancing their grades.



Figure 5.2. Studio physics setup.

Credits: TEAL project MIT

Assessment issues. The current system of feedback, and formative assessment, is also, not geared, to assess an important aspect of student learning, which is meant to be the primary target of formative assessments - what students know about the concepts before, and what they learned about the concepts after each experiment (NRC 2001). The Pre-lab quizzes, could be expected to provide information, about learners' prior constructions of concepts related to the inquiries conducted during studio labs. But currently, they are made up of multiple choice questions, and like Clickers, do not yield tangible information about the substance of students' thinking. Per Vygotsky (1986), learners' articulations symbolize their thoughts, and/or precede their thoughts. Thus, to obtain substantive information about learners' thinking process, it is necessary to have them articulate their reasoning process, as was done through PreQ and PostQ, in this study. Although Labs do provide partial information about learners' thinking after the experiments, they are not interpreted to extract such information, as was done in this study, by using ENA to interpret PreQ and PostQ. Valuable opportunities of formative assessment, are thus being lost, because of the way these three assessment instruments, HW Quiz, Labs, and Lab Quiz, are currently structured, delivered, and/positioned. And since formative assessments are a vital part of active learning, students and TAs are missing opportunities to gauge learning, and therefore, to redirect students' misconceptions, and/missed connections as identified through this study.

Evidence. Some of the unresolved knowledge and/ belief formulations, which emerged through this study, were as follows:

(a) Some students seemed to have retained the naïve expectation, that ‘electric shock’ was supposed to be an effect of holding a voltage source, and/ probes of a multimeter with wet hands (Experiment 1)

(b) Many students missed making connections to area of a parallel plate capacitor, being instrumental in changing its capacitance, along with distance between the plates (Experiment 2).

(c) Many students missed making connections to their mathematical competencies of interpreting and/discriminating between linear and non-ohmic graphs, which led them to make incorrect conclusions about LED and light bulb being ohmic materials (Experiment 3).

(d) Students demonstrated intuitive formulations about the behavior of current in RC circuits as ‘wave-like’, for instance. (Experiment 4)

(e) Students did not connect to relevant canonical knowledge, and articulations, for making appropriate inferences about relationships between physical quantities, based on their inquiry, e.g. capacitance, and area; behaviors of voltage and current in RC circuits; the graphical output of Voltage-current and its relationship to the nature of resistance (ohmic or non-ohmic).

(f) Strong intuitive and transitional connections instead of canonical connections, resulted in negative sophistication gains of experiments 3 and 4.

Summary. The findings of this study, thus point to the importance of capturing the relative strengths of all types of connections, to evaluate sophistication of thinking. It is clear, that the epistemological issues in thinking cited above, would have remained uncovered, were they not captured by the ENA-based assessment implemented during studio 4. Even now, they have only been captured, but remain unaddressed. Although research-based active learning environments are well-integrated into course design, the assessment components were not meaningfully formative, for the following reasons:

(a) they were not strategically situated in the delivery of instruction (as is the case with lab Quiz),
(b) they did not integrate appropriate technology for delivery (Lab Quiz, Labs),
(c) the observations were not interpreted to reveal substantive information on student thinking (as is often the case with lab reports and quizzes),

(d) assessment items were not designed to capture learners’ reasoning process explicitly (as is the case with Pre-lab quiz).

Significance. Because students of most non-Majors introductory physics courses, typically do not take advanced physics courses, it is alarming to think, that they could be going away with a semester full of unclear formulations about basic concepts in electricity, magnetism, and modern physics, which have direct applications in the workings of all minor and major electrical and electronic gadgets we use, in every-day life, from a battery, to a cell phone, or the laptop. Apart from the loss of valuable learning opportunities, these ‘unwarranted takeaways’ by learners, from physics courses, add to the negativity of

perceptions surrounding the study of physics in society at large, which by itself, should be a cause for concern, since participation in physics has shown very little improvement, in comparison to other natural sciences, despite significant advances made by physics education research, in the last 50 years, in improving our understanding of how students learn physics, and in the adoption of research-based, student centered, instruction and curricula.

Recommendations. I contend that, paying attention to two components of instruction design could be significantly useful in resolving the issues summarized above– (a) strong alignment between learning goals, instruction, technology, and assessment, which is also recommended strongly in most NRC reports, but less researched in PER, and (b) emphasis on the formative, rather than the summative aspect of assessment, such that connections in student thinking are made visible and interpreted, as was accomplished through the ENA-based assessment implemented in this study (NRC 2001). Research suggests that the following principles should guide the development of formative assessments in inquiry contexts, which seek to prioritize learning over assessment: (i) Obtain information about learners’ prior constructions about the concepts being inquired; (ii) have students describe their concrete-experience of observations and findings (iii) have students construct meaning, make inferences and connect explicitly, to canonical knowledge (iv) have students make real life connections (Minstrell, Anderson, and Li, 2011).

The question prompts in this study, adopted a similar approach, as recommended above, by asking 3 questions for each experiment – 1 PreQ, and 2 PostQ, in which one PostQ asked students to describe findings, while the other PostQ asked them to either justify, or articulate their inferences, and most students were unable to use canonical expressions to describe or infer relevant relationships. While existing Lab reports do ask students to answer questions having objective 2, and sometimes, objective 3, objectives 1 and 4 remain unaddressed. The following modifications to Labs, HW Quiz, and Lab Quiz would enhance alignment, and formative aspect, two criteria identified above, as significant for enhancing student learning.

Pre-Lab Quiz. With every multiple-choice question, include an open-ended question (which need not be included in grading) – a text entry box, which asks students to explain their reasoning. The responses could be either quickly reviewed manually by the TAs, or instructor, to get a sense of students’ thinking prior to the Labs, or some automation could be developed by using text analytic software, having word-cloud feature, to do the analysis. Further analyses informed by ENA could be utilized to inform instruction (content development), refine questions, and advance research. One other component which could be introduced here is the ‘muddiest points’ reflections which have gained popularity with many instructors. So, towards the end of the Quiz, or after specific questions which are known to have a higher difficulty index, a question prompt asks students what concept, or idea was ‘muddiest’ or unclear, or ambiguous to them, and why. When the ‘why’ question is asked, student thinking, and especially their

epistemic framing, becomes more visible. This also makes it easy to know ahead of class, where the class thinking stands with respect to specific topics.

Lab Quiz. Replace the Lab Quiz by PreQ, asking conceptual and/problem questions related to the experiment, which explicitly ask students to explain their answer, and/ give reasons. Since each experiment, would be preceded by a PreQ, the number of questions would remain the same as those on existing quizzes. It could be administered via Twitter or using Canvas. The benefit of using Twitter, would be, that data can be extracted almost immediately onto spreadsheet, for quick preliminary analysis by the TAs. Spreadsheet download also enables the immediate use of Text analytic software to generate word clouds, for instance. This will allow the TAs to get a preview into learners' prior constructions and make clarifications during activities. The other option is to use Clickers (preferably text-enabled) instead, to administer the Quizzes.

Labs. Re-Structure the lab report questions, so that each experiment asks students to describe-infer-connect to real world; equivalent to 3 PostQs, of ENA-based assessment. Replace the submission of hard copies by soft copies on Canvas. For graphical outputs, screenshots, or saving them to pdf rather than printing, could be explored. It will not only save printing paper but enable TAs to download the labs directly on to their computer, comment, grade and re-send to the students for revisions if they wish to do so, during next labs. It will also enable TAs to pull up the content which relates to common errors, onto the projection screen, and discuss or clarify common mistakes to students, together. A group discussion becomes possible.

Lab feedback. Use any Canvas feature, such as Discussion boards, or chat, or Class Twitter account, to enable virtual classroom dialogue between students, about questions, doubts, etc. related to the activities, which can also be projected onto the projection screen for everyone to see the flow of conversation, so that the option of knowledge construction through social interaction is not limited to student-TA interaction, but peer to peer exchanges between groups are also enabled. TAs can use these forums to clarify doubts, give common feedback on their observations of students' observations. Specific participation points, and/extra credit, for engaging actively in dialogue could be assigned, based on the number and relevance of questions, and answers exchanged.

More Avenues in Practice

Apart from the studio lab, QOT2SM assessments could be realized in other learning contexts, and other undergraduate physics courses, including Majors physics, where complex thinking is normally involved, in the sense that learners' cognitive elements comprise of multiple representations of knowledge – mathematical, graphical, schematic, and where spatial thinking capabilities play a significant role in

student performance. In these learning scenarios, it becomes vital, that student thinking is made visible to the instructor, for ensuring learning gains, and equity; because learner differences in thinking have been related to differences in performance, and participation in physics. Some examples of potential usage are outlined below.

Large-lecture Sessions. While Twitter was used in this study, for its back-end compatibility with other applications necessary for qualitative, and quantitative analyses, the QOT component, in a large lecture setting, could be implemented using alternative technologies – stand-alone devices such as text-enabled Clickers, or features of course management systems (CMS) such as discussion boards, and chat boards, which can also, like Twitter, and Clickers be projected on to the screen, so that students are not only engaged in classroom discourse, but the instructor can use the projected conversation, to redirect learners' attention towards 'missing' elements of knowledge, to address misconceptions, and/clarity ambiguous constructions demonstrated by learners' responses. The prompts could be specific, to gain information on understanding, or could also, like the 'muddiest points' reflections, seek information on 'difficulties' or 'perceptions', or 'prior beliefs', 'expectations', and so on.

For that matter, any block of the QOT2SM assessment framework, could be customized to suit the needs of instruction, and could be aligned with learning goals. It is true, however, that QOT cannot be used to generate an instantaneous aggregated graphical output on the lines of Clickers, unless suitable customization of currently available Web versions of open-source text-analytic packages, is carried out, or open-source environments such as Python, which are text-friendly, are used to generate customized scripts, to identify SEs, like RScript was developed in this study, to compute SMs. The information can still, however, be processed, subsequently to generate a course catalogue of 'student difficulties', or be used to modify instruction, or inform research.

Smaller upper division Majors physics courses. Since ENA was basically developed to enable statistical modeling of cognitive networks or behavior and/performance patterns in complex problem-solving situations (Shaffer, 2017), it can also be utilized to identify trends, barriers, and patterns of connections in learners' thinking, in upper division physics classes. While Epistemic framing analyses have previously been conducted in these contexts in physics, to analyze students' problem-solving behavior, and/ their use of math in physics, ENA has not been tried, in these contexts, either. QOT in these contexts, can be video observations, and/student generated explanations of their thinking, triangulated with regular artefacts generated through routine course assessments. The analyses will not only be useful for informing instruction, but also inform research in learner differences in thinking, expert-novice differences, and advance knowledge about how learners' thinking shifts with the use of

different instructional aids in upper division classes, or even to make comparisons between classes, or to evaluate instruction.

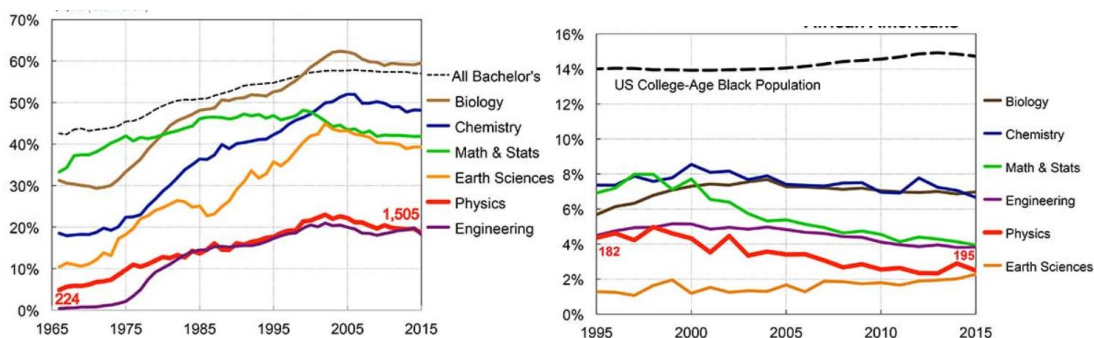
Virtual Learning-Assistant. Learning Assistant programs, LAs as they are commonly referred to, are implemented to enhance student learning, through peer-to-peer interactions. However, most of these programs, including the Tutoring Program at the Department of Physics, at UTK, are implemented in physical spaces, where students meet, and hang out in informal settings, and work through problem sets, using creative aids. While these programs have yielded positive results (Sellami, et.al., 2017), the interactions between LAs and students, which could be rich sources of data, to inform instruction and research, are lost, because of the platform chosen for implementation. These could as easily be implemented online, to enhance - flexibility for the LAs and students, participation, interactions, and therefore, social construction of knowledge. Each course website is a good platform for scheduling asynchronous (like the homework help available online) discussions/sessions, or synchronous sessions by consensus. The important thing is, every digital interaction has a provision for being recorded, which allows instructors to keep track of, check in, and analyze the rich data available, by selecting specific segments, to inform instruction and research.

Implications for Research

The findings of this study are significant for addressing three specific research problems – (i) Inequity of access to physics (ii) Bridging the gap between educational neuro-science and learning science (iii) Real-time Utilization of QOT2SM for formative assessment. I will discuss each of these issues in the following paragraphs, by drawing upon relevant literature, to support my argument, and outline avenues of research, which should be pursued.

Enhancing Equity of Access to Physics

By the latest statistics available, enrollment at the University of Tennessee, in Fall 2016, by department, and major, was reported as 86% male to 14% female in physics; 63% male to 37% female in math; 80% male to 20% female in engineering; and the male to female ratio for Chemistry and Biology, was 54% to 45%, and 48% to 52%, respectively (Online reports @ oira.utk.edu). The underrepresentation of women (Figure 5.3) and individuals from minority segments of society (Figure 5.4) in physics, has been an issue, which has perplexed educators and policy-makers for many decades. Gender inequity in physics, has been an issue close to heart for me, ever since I enrolled in the LEEDS Doctoral program, and became aware of it, through coursework. So, it was also the initial focus of my Dissertation.



Credit: APS/Source: IPEDS Completion Survey

However, as I reviewed literature, and conducted small qualitative studies, to examine gender differences in classroom interactions which happen in social sciences, versus physical sciences; and a pilot study, with the same population, as this study, to examine gender differences in perception of inquiry activities, I realized, that there was a need for finding ways to collect large sample qualitative data on thinking in-context, and ways to analyze the data qualitatively and quantitatively. Because research pointed towards three prime candidates as being instrumental in contributing to the gender gap – Spatial reasoning (Hedges & Nowell, 1995; Kimura, 2002; Halpern, Aronson, Reimer, Simpkins, Star, & Wentzel, 2007), cognitive styles (Severiens & Geert, 1994; Stadler, Duit, & Benke, 2000), and epistemological beliefs Belenky, et al. (1986, 1997); but there seemed to be two problems with data collection and analyses in conducting valid gender studies – one, quantitative data was inadequate to map learners’ thinking process, and therefore, their reasoning and framing, and two, although qualitative data presented a good fit, the analytical techniques constrained sample size, raising issues of validity of results. So, I began to research how best to gather qualitative data for large samples (necessary for generalization), and how to quantize it without losing its richness. This quest, led me to Epistemic Network Analysis, and to the conceptualization of this study.

Thus, while I am aware that the assessment instrument developed for this study, needs further validation, and refinement, I know that large sample gender studies based on QOT, will indeed be possible, with the proposed framework. While gender was initially included as a factor in this study, I did not pursue it further, since the entire sample of 38 participants, consisted of only four male participants. I have had the same problem while investigating gender as a factor in a social sciences course, and in a high school robotics club. The class had majority women and only five male students; the high school robotics club ‘build’ team versus ‘programming’ and ‘business’ team investigation, also had a dominance of males on the ‘build’ team, and females on the ‘business’ and ‘programming’ teams.

Another striking problem with the two times that I conducted studies in the present context (PHYS 222), is, that females were more willing to participate in research studies than males. This has been the trend with the robotics club, and social sciences class as well. So, knowing that in PER, educators are willing to explore innovative, research-based instructional and curricular strategies, I hope that formative assessments, as proposed in this chapter, and the study, will also be received with openness, especially since it is evident from the findings, that very different, and concrete information about connections in thinking, and therefore, the thinking process itself, becomes available by assessing SNs and SMs, than is made available, by assessing only for SEs, as is done by all existing assessments. Such classroom data can be used for identifying and documenting learner and gender differences in thinking, and thereby, make a valuable contribution towards reducing inequity in physics.

Correlating Neural Networks and Cognitive Networks

Spatial thinking ability – which comprises of three distinct skills, spatial perception, mental rotation, and spatial visualization, has been shown to be a reliable predictor of success in physics (Kozhevnikov, Hegarty, & Mayer, 2002; Kozhevnikov, & Thornton, 2006). Also, several fMRI (functional magnetic resonance imaging) studies (Weiss, Kemmler, Deisenhammer, Fleischhacker, & Delazer, 2003; Unterrainer, Wranek, Staffen, Gruber, & Ladurner, 2000; Thomsen, Hugdahl, Ersland, Barndon, Lundervold, Smievoll..., Roscher, & Sundberg, 2000), show evidence of sex differences in brain activation while performing mental rotation tasks, with males outperforming females, in significant number of studies. Different regions of the brain were activated in males and females, and differences in activation levels were found, during various stages of the mental rotation task, with men using object rotation/recognition while women used a more analytic/serial strategy in the task. These two separate findings, one from neuroscience, and the other, about spatial ability from the Learning sciences, sparked interest in investigation of neural correlates of cognitive processes underlying spatial reasoning. While there is still no consistency between findings for drawing conclusions about spatial abilities, neuro-imaging studies began to receive greater interest in PER. A neuro-imaging study based on pre-post fMRI scans of brain activity as students solved a question from the force concept inventory, found “increased brain activity notably in lateral prefrontal and parietal cortices that previously have been associated with attention, working memory, and problem solving, and are collectively referred to as the central executive network” (Brewer, et. al., 2018).

The application of neuroscientific studies about human learning and cognition, to educational practice has been received in the realms of education sciences with a lot of skepticism, since brain-based learning first made its appearance in this field, in the 90s. Many studies later, in both fields, it now makes perfect sense, that cognitive scientists (or learning scientists for that matter), and neuroscientists, who are

researching the same subject of human cognition and learning, by studying - how the human brain processes information, and learns higher order thinking skills; and how the anatomy and physiology of the brain correlate to human cognition and learning, respectively, should be using research from the other field, to inform their own studies. Exploration of data collection and/ analytical techniques, which are shared by both fields, could probably accelerate the process of ‘sharing and exchanging knowledge’, since for instance, it will facilitate the conduct of correlational studies, which examine how students process specific problem-solving tasks in physics, differences in experts and novices, and gender differences. Recently, eye-tracking technology, was used to collect data about underlying cognitive processes, as alternative to self-report techniques and think-aloud interviews. Gaze patterns were used as proxies to identify patterns in cognitive function (Was, Sansosti, and Morris, 2016).

Epistemic network analysis has previously been used in analyzing neural networks, gaze patterns, as well as cognitive networks in the context of complex thinking (Shaffer, 2017). Thus, advancing the usage of QOT2SM assessments, as implemented through this study, in undergraduate physics courses, is significant on two accounts – (a) it presents instructors an analytical technique which integrates three modalities flexibly, to assess sophistication of thinking – qualitative analyses to assess SEs, network analyses to assess SNs, and quantitative analyses to assess SMs; and (b) it presents education researchers a technique which can be used to correlate cognitive networks of learners realized through QOT, with their neural networks, and/ their eye-gaze patterns, during performance of the same domain-specific task. For instance, the Brewe, et.al. (2018) study data could be correlated with the students’ self-reported reasoning for solving the FCI question, the way he/she did, and it could also be correlated with student’s eye-movement data, using ENA. Such triangulation studies, will not only serve the purpose of validation of research and theory, in the Learning sciences, but advance the development of cross-disciplinary technologies, which could radically change the ways we use them in education practice. I found this quote from the editorial page of the open access journal, ‘Trends in Neuroscience and Education’, very appropriate to make this point.

“**Neuroscience** is to education what biology is to medicine and physics is to architecture. Biochemistry is not enough to cure a patient, and physics is not enough to build a bridge. But you cannot perform great work, neither in medicine nor in architecture, against the laws of physics or biology. And in fact, they will inform you about many constraints and rule out a great many of projects right from the start as failures.”

Technology Integration for Automated Execution

As discussed earlier, the QOT2SM framework is open and flexible, and alternative technologies to collect qualitative observations of thinking, can be integrated. The primary criterion of technology choice in every block, must be, that it has good interface compatibility with the technology used in the

next block. However, I will discuss below, two technologies, which present a good fit for realizing seamless automation and execution, and which I recommend, be explored – (i) nCoder (ii) rENA.

nCoder. nCoder is a Text analytical tool, developed by the researchers of Epistemic Analytics group, at the University of Wisconsin-Madison, to develop, validate, and implement, automated coding of large sample textual data (<http://n-coder.org/>), like the self-reported qualitative observations of student thinking, derived in this study, using Twitter. It can be utilized for qualitative coding of any form of textual data – to minimize the amount of data needed to be coded manually, using statistical techniques to establish validity and reliability, and to apply the validated coding scheme, automatically, to the larger dataset. Raw data organized into columns is uploaded as csv file, utterance column is chosen, and code list created by defining code descriptions and adding expressions (words, and/phrases) to be coded. It is then validated, by calculating Cohen's kappa, and tested and applied to the remaining data. It has been used previously, to code transcripts of conversations and interactions between surgeons during surgery, and online chat conversations and log files of epistemic games, and virtual internships (<http://n-coder.org/>).

While other text analytic packages on the web could also be used for the same purpose, the advantage of using nCoder, just like ENA Web Tool, is that it was developed specifically for coding educational data, and therefore, user-friendly from the perspective of educators, who have minimal background in qualitative coding, and/software programming. The output generated is compatible with the formatting requirements of ENA Web Tool, the technology, which presumably will be used in the next block (SN).

rENA. rENA is a package available in the RStudio environment (<https://CRAN.R-project.org/package=rENA>) and can be used to generate ENA sets (rData) from the coded csv data of QOT, which contains SEs, and their occurrences, along with metadata. It can be used to plot network graphs, conduct t-tests, and extract them as outputs, a feature (extraction), which is not available on the ENA Web Tool. Thus, it could present an efficient way to conduct network and quantitative analyses using the same technology RStudio and eliminate the hassle of exporting ENA sets to RStudio, thereby enhancing automated execution, especially since RScripts can be written to the instructor's specific requirements of network, and quantitative outputs desired. rENA was also developed by the same research group who developed ENA Web Tool and nCoder. I expect that the two technologies presented here, would accelerate the assessment process, and take the QOT2SM framework much closer to implementation of real-time formative assessments of sophistication of thinking, in studio settings, and it is the next step of my research plans.

I will close this chapter, and thesis, with my concluding thoughts on why I think this line of research has been missing from PER for quite some time, why I believe it should be pursued urgently, how the departments of physics need to open their doors to ‘outsiders’ to further this research, and how University of Tennessee is placed uniquely, to take an initiative in this direction.

Concluding Thoughts: Technology and Assessment

The integration of technology into curricula and instruction, in Undergraduate Physics education, has been driven by a valid need to enable visualization of abstract concepts, and to engage students actively in knowledge construction, through inquiry and/ problem solving activities (McDermott, 2001). Wide ranging reform-based curricula and student-centered instructional environments, rich in technology, have been deployed across physics departments, and universities in the US, because these approaches have been found to yield significant learning gains, as opposed to traditional instruction-centered approaches, which emphasize lecture delivery (Sharma, Johnston, Johnston, Varvell, Robertson, Hopkins, Cooper, & Thornton, 2010; Hake, 1998). Most of these ALEs (active learning environments), are developed by one or more physics departments and /universities, who focus on physics education research, which seeks to understand how students learn physics, and use that research, to develop specialized instructional environments, such as studio physics (Meltzer, and Thornton, 2012). While these techniques are being widely utilized by faculty teaching undergraduate courses, for enrichment, integration of technology, and design of formative assessments, are often not given the careful attention and consideration required, to make them effective in sophisticating learners’ thinking (NRC 2001).

Alignment and Commitment

NRC reports over the years, are found to be replete with guidelines and frameworks for achieving alignment between learning goals, instruction, choice of technology, and assessment – which are the fundamental principles (building blocks) of any model of Systematic Instruction Design (NRC 2001). However, through my observations as an educator, and as a researcher, I have gathered that two factors stand in the way of implementing these guidelines in practice: (i) physics educators and many a times, physics education researchers, at the university level, do not quite grasp the significance of such alignment, possibly because they have no theoretical knowledge or training in pedagogy, and/ instruction design, and (ii) too often, other research and academic commitments in physics as such, prevent faculty as well as the TAs, from investing the time needed, for continuous monitoring, and realignment of instruction and curricula, based on feedback received from assessments. At the institutional level, it could also be, that physics departments have been enthusiastic about adoption of these specialized curricula and ALEs but have either not recognized the significance of systematic instruction design, which seems

unlikely, or, are lacking in faculty, who might be willing and/ committed to invest the time necessary, for effective implementation.

Technology in assessment. As I have discussed throughout this thesis, the other big problem, apart from lack of attention to systematic instruction design, is, that integration of technology for regular formative assessments in the classroom, has not received the same attention in PER, as has been given to its integration into instruction, for enhancing student thinking and learning. So, while there is recognition of the need to make student thinking visible, and much research has been done to find ways to make student thinking visible, there is indeed a gap in research, in the deployment of technology for designing systematic formative assessments aligned with instructional activities, and their goals. As I say this, I realize that some would point to Intelligent Tutors, and other individualized assessments systems built into these technologies, to build upon student thinking. However, the fact remains, that these systems seem to be years away (to be lenient), from being routinely adopted for classroom usage.

Urgency. In the meantime, as the findings of this study suggest, in a limited context no doubt, that students may be leaving undergraduate physics classrooms with partially intuitive constructions of the concepts they were taught, using one of the most researched, well-documented, and widely used ALEs – studio physics, simply because existing assessments fail to capture these epistemological issues. As I have discussed in the previous sections, this reason, should be more than enough cause for concern, given that (a) people have negative attitudes towards participation in physics, and (b) participation rates in physics have been significantly low for decades, despite our best efforts at reform.

QOT2SM. The adoption of QOT2SM framework, for designing assessments, could go a long way towards capturing and resolving learners' misconstructions, and enabling canonical constructions, since students and TAs will get opportunities to interact, and give feedback in-context. An equally important reason for its adoption, is its ability to capture learner differences in cognitive processing and framing, quantitatively, without losing the qualitative substance in data, in-context, which makes it a good assessment candidate, as an analytical tool for conducting large sample gender studies, vitally important, for enhancing gender equity, and equity in general, in physics.

Enhancing Systematic Instruction Design

The problems highlighted above, speak of a lack of systematic instruction design, which includes the integration of technology in the structuring of curricula, instruction, and assessments in physics courses. In view of the severely negative implications of this deficiency, for equity and participation in physics, I think it is the moral and ethical responsibility of every department of physics, across the country, to take affirmative action to overcome this deficiency, by implementing appropriate policies,

which enable and promote best practices of technology-enhanced instruction design. I propose, here, three tentative solutions: TA Training, Education Technology and Physics Specialist, PER graduate program.

TA Training. TAs are the backbone of undergraduate physics courses, whether the course adopts studio physics, labs and recitations, or any other reform-based approaches to design. Success and effectiveness of the course rests heavily on their shoulders. While they are committed to physics, they are either not invested in paying attention to systematic instruction design, or are not equipped to do so, or both. This problem can be addressed by having them complete either online or face-to-face certificate programs specially designed to include 3 to 5 courses each on ‘learning and assessment’, ‘PER’, ‘Technology integration’, and ‘Action research 1’, ‘Action Research 2’. Action research could be the culminating component of the program, which will allow them to implement interventions in their own courses and study its impact. A ‘teaching certificate’, and/course credits could be used as incentives to complete all courses, while a minimum of two courses, could be made mandatory for continuance of teaching assistantships.

Instruction Technology and Physics specialist. The second solution, would be to recruit full time specialist/faculty, who specializes in instructional technology and physics education research, or allied fields, to coordinate and supervise technology integration, in various courses, through one-on-one consultation with faculty. Although, it might be argued that OIT is equipped to meet this need, I would counter, that teaching and learning physics is a specialized enterprise, and a dedicated specialist as described above, would be needed for success of technology integration, in a specialized domain.

PER as an area of study. If physics education research, is integrated as an area of study, at the graduate level, it will likely foster a culture of awareness about its significance, and of systematic, research-based alignment of goals, instruction, technology, and assessment. This will also incentivize existing faculty, TAs, and graduate students, to pursue research in this area, and contribute their time and effort. Opportunities will open for recruitment of scholars from Education Sciences, who are interested and committed to the cause of PER, and the interdisciplinary exchanges facilitated through such recruitments, will make significant positive contribution towards promoting a culture of treating physics education as a field, at par with other areas of research in physics.

Online programs. Rates of participation, in these graduate programs can be significantly enhanced, by adopting a model of ‘online Masters in physics, and in physics education’ on the lines of Texas State University, which allows high school physics teachers to enroll in these programs, while working, thereby opening physics department to the community at large, while also enriching the departmental culture, through exchange of pedagogic ideas with trained, practicing teacher-students.

University of Tennessee. While many models of programs in PER already exist across the country, there are many gaps by geography, and the state of Tennessee, occupies one such gap. It should be noted here, that the University of Tennessee, is favorably placed to fill this gap, on two counts. One, it has an excellent program in Education, which was instrumental, in the state being a recipient of the federal grant, ‘race to the top’. Two, the college of education offers a wide range of programs in Learning sciences, STEM Education, Instructional Technology, Educational Psychology, and Ed Policy, which could be tapped for collaborative ventures, by the physics department, to initiate the said program.

Next Steps

As I consider my next steps, I know that for me, an opportunity, to take on the challenge of working as an education technology specialist dedicated to the cause of evaluating and reforming curricular, instructional, and assessment design in a physics department, would always get first preference over other courses of action and research avenues, I plan to pursue. I am currently exploring ways by which I could build upon this Dissertation research. At the top of my list of options to explore, are –

(a) Reanalyzing the existing data, and the data from studio 3, with nCoder and rENA, to refine and accelerate the assessment process

(b) Collaborating with existing physics faculty, to implement the assessment as part of their course, in large -lecture setting, or in other Majors undergraduate physics courses,

(c) Using the QOT2SM framework to conduct gender studies in undergraduate physics

(d) Applying for NSF and/Spencer grants for conducting my proposed studies.

(e) Knowledge building, by submitting to Journals and Conferences – AAPT, PERC, NARST, and AECT.

LIST OF REFERENCES

Chapter 1: Introduction

- Arastoopour, Chesler, D'Angelo, Shaffer, Opgenorth, Reardan, Haggerty, & Lepak. (2012, June). Nephrotex: Measuring first-year students' ways of professional thinking in a virtual internship. Paper presented at American Society of Engineering Education (ASEE), San Antonio, TX.
- Bao, L., and Redish, E. (2006). Model analysis: Representing and assessing the dynamics of student learning, *Phys. Rev. ST Phys. Educ. Res.* 2, 010103.
- Beckett, & Shaffer. (2005). Augmented by reality: The pedagogical praxis of urban planning as a pathway to ecological thinking. *Journal of Educational Computing Research* 33 (1): 31–52.
- Bing, T. J. (2008). An epistemic framing analysis of upper level physics students' use of mathematics. ProQuest.
- Bodin, M. (2012). Mapping university students' epistemic framing of computational physics using network analysis. *Physical Review Special Topics-Physics Education Research*, 8(1), 010115.
- Cowie, B. (2012). Focusing on the classroom: Assessment for learning. In *Second international handbook of science education* (pp. 679-690). Springer, Dordrecht.
- Cummings, K. (2011). A developmental history of physics education research. In *Second Committee Meeting on the Status, Contributions, and Future Directions of Discipline-Based Education Research*. Available: http://www7.nationalacademies.org/bose/DBER_Cummings_October_Paper.pdf.
- Docktor, J. L., & Mestre, J. P. (2014). Synthesis of discipline-based education research in physics. *Physical Review Special Topics-Physics Education Research*, 10(2), 020119.
- Donovan, M. S., & Bransford, J. D. (Eds.). (2004). *How students learn: science in the classroom*. Retrieved from <https://ebookcentral-proquest-com.proxy.lib.utk.edu:2050>
- Elby, A., & Hammer, D. (2001). On the substance of a sophisticated epistemology. *Science Education*, 85(5), 554-567.
- Epistemic Network Analysis. (2017, April 4). Retrieved from <http://edgaps.org/gaps/projects/epistemic-network-analysis/>.
- Fredlund, T. (2015). Using a Social Semiotic Perspective to Inform the Teaching and Learning of Physics.
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American journal of Physics*, 66(1), 64-74.

- Hammer, D. (1994). Epistemological beliefs in introductory physics. *Cognition and Instruction*, 12, 151-183.
- Hammer, D., & Elby, A. (2003). Tapping epistemological resources for learning physics. *The Journal of the Learning Sciences*, 12(1), 53-90.
- Hill, C., Corbett, C., & St Rose, A. (2010). Why so few? Women in science, technology, engineering, and mathematics. American Association of University Women. 1111 Sixteenth Street NW, Washington, DC 20036.
- Hofer, B. K. (2004). Exploring the dimensions of personal epistemology in differing classroom contexts: Student interpretations during the first year of college. *Contemporary Educational Psychology*, 29(2), 129-163.
- Junco, Elavsky, & Heiberger. (2013). Putting twitter to the test: Assessing outcomes for student collaboration, engagement and success. *British Journal of Educational Technology*, 44(2), 273-287.
- Keiner, L. E., & Burns, T. E. (2010). Interactive engagement: How much is enough? *The Physics Teacher*, 48(2), 108-111.
- Kober, N., Board on Science Education, & National Research Council of the National Academies. (2015). Reaching students: What research says about effective instruction in undergraduate science and engineering.
- Kohl, P. B. (2007). Towards an understanding of how students use representations in physics problem solving. ProQuest.
- Kozma, R.B., The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction*, 2003. 13: p. 205.
- Larkin, T. L. (2011). Assessment of student understanding in physics: An integrated qualitative and quantitative approach. In American Society for Engineering Education. American Society for Engineering Education.
- Laws, Sokoloff, & Thornton. (1999). Promoting active learning using the results of physics education research. *UniServe Science News*, 13, 14-19.
- Lising, L., & Elby, A. (2005). The impact of epistemology on learning: A case study from introductory physics. *American Journal of Physics*, 73(4), 372-382.

- May, D. B., & Etkina, E. (2002). College physics students' epistemological self-reflection and its relationship to conceptual learning. *American Journal of Physics*, 70(12), 1249-1258.
- McDermott, L., A view from physics, in *Toward a scientific practice of science education*, M. Gardner, et al., Editors. 2004, IOS: Amsterdam.
- McDermott, (2001). Oersted medal lecture 2001: "Physics Education Research—the key to student learning". *American Journal of Physics*, 69(11), 1127-1137.
- McDermott, L. C., & Redish, E. F. (1999). Resource letter: PER-1: Physics education research. *American journal of physics*, 67(9), 755-767.
- Meltzer, & Thornton. (2012). Resource letter ALIP-1: active-learning instruction in physics. *American journal of physics*, 80(6), 478-496.
- Michael, J. (2006). Where's the evidence that active learning works? *Advances in physiology education*, 30(4), 159-167.
- Nash, & Shaffer. (2008). Player-mentor interactions in an epistemic game: A preliminary analysis. Paper presented at the International Conference of the Learning Sciences (ICLS), Utrecht, Netherlands.
- National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. National Academies Press.
- Nulty, & Shaffer. (2008). Digital Zoo: The effects of mentoring on young engineers. Paper presented at the International Conference of the Learning Sciences (ICLS), Utrecht, Netherlands.
- Redish, E. F., & Steinberg, R. N. (1999). *Teaching Physics: Figuring Out What Works*.
- Redish, E. F (2004). A theoretical framework for physics education research: Modeling student thinking. arXiv preprint physics/0411149.
- Robbins, S. (2011, March). Beyond clickers: using ClassQue for multidimensional electronic classroom interaction. In *Proceedings of the 42nd ACM technical symposium on Computer science education* (pp. 661-666). ACM.
- Scherr, R. E., & Hammer, D. (2009). Student behavior and epistemological framing: Examples from collaborative active-learning activities in physics. *Cognition and Instruction*, 27(2), 147-174.
- Shaffer, (2006). Epistemic frames for epistemic games. *Computers & Education*, 46(3), 223-234.
- Shaffer, & Graesser. (2010). Using quantitative model of participation in a community of practice to direct automated mentoring in an ill-formed domain. Paper presented at *Intelligent Tutoring Systems (ITS)*, Pittsburgh, PA.

- Shaffer, Hatfield, Svarovsky, Nash, Nulty, Bagley, Frank, Rupp, & Mislevy, (2009). Epistemic Network Analysis: A prototype for 21st Century assessment of learning. *The International Journal of Learning and Media*. 1(2), pp. 33-53.
- Shaffer, & Ruis. (In press) (2017). Epistemic network analysis: A worked example of theory-based learning analytics. *Handbook of Learning Analytics and Educational Data Mining*.
- Sharma, M. D., Johnston, I. D., Johnston, H., Varvell, K., Robertson, G., Hopkins, A., Stewart, C., Cooper, I., & Thornton, R. (2010). Use of interactive lecture demonstrations: A ten-year study. *Physical Review Special Topics-Physics Education Research*, 6(2), 020119.
- Singer, S. R., Nielsen, N. R., & Schweingruber, H.A. (2012). *Discipline-Based Education Research*. Washington: National Academies Press. Retrieved from **<http://ebookcentral.proquest.com.proxy.lib.utk.edu:90/lib/utk/detail.action?docID=3379024>**
- Svarovsky, & Shaffer. (2007). SodaConstructing knowledge through exploratoids. *Journal of Research in Science Teaching* 44 (1): 133–153.
- Tuminaro, J., & Redish, E. F. (2007). Elements of a cognitive model of physics problem solving: Epistemic games. *Physical Review Special Topics-Physics Education Research*, 3(2), 020101.
- Yerdelen-Damar, S., Elby, A., & Eryilmaz, A. (2012). Applying beliefs and resources frameworks to the psychometric analyses of an epistemology survey. *Physical Review Special Topics-Physics Education Research*, 8(1), 010104.

Chapter 2: Literature Review

- Arastoopour, Shaffer, Swiecki, Ruis, & Chesler. (2016). Teaching and assessing engineering design thinking with virtual internships and epistemic network analysis. *International Journal of Engineering Education*, 32(2).
- Bao, L., and Redish, E. (2006). Model analysis: Representing and assessing the dynamics of student learning, *Phys. Rev. ST Phys. Educ. Res.* 2, 010103.
- Bateson, G. (1972). *Steps to an ecology of mind: Collected essays in anthropology, psychiatry, evolution, and epistemology*. University of Chicago Press.
- Baxter Magolda, M. B. (1987). The affective dimension of learning: Faculty-student relationships that enhance intellectual development. *College Student Journal*, 21, 46-58.

- Baxter Magolda, M. B. (1992). *Knowing and reasoning in college: Gender-related patterns in students' intellectual development*. San Francisco: Jossey Bass.
- Bing, T. J. (2008). *An epistemic framing analysis of upper level physics students' use of mathematics*. ProQuest.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. National Academy Press.
- Belenky, M. R., Clinchy, B. M., Goldberger, N. R., & Tarule, J. M. (1986). *Women's ways of knowing: The development of self, voice and mind*. New York: Basic Books.
- Bodin, M. (2012). Mapping university students' epistemic framing of computational physics using network analysis. *Physical Review Special Topics-Physics Education Research*, 8(1), 010115.
- Buehl, M. M., Alexander, P. A., & Murphy, P. K. (2002). Beliefs about schooled knowledge: Domain specific or domain general? *Contemporary Educational Psychology*, 27, 415–449.
- Carey, S. (1999). Sources of conceptual change. In E. K. Scholnick, K. Nelson, & P. Miller (Eds.), *Conceptual development: Piaget's legacy* (pp. 293–326). Mahwah, NJ: Erlbaum.
- Carey, S. (2000). Science education as conceptual change. *Journal of Applied Developmental Psychology*, 21, 13–19.
- Carey, S., & Smith, C. (1993). On understanding the nature of scientific knowledge. *Educational Psychologist*, 28(3), 235-251.
- Caramazza, A., McCloskey, M., & Green, B. (1981). Naive beliefs in “sophisticated” subjects: Misconceptions about trajectories of objects. *Cognition*, 9(2), 117-123.
- Chinn, C. A., Buckland, L. A., & Samarapungavan, A. L. A. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology. *Educational Psychologist*, 46(3), 141-167.
- Collins, A., & Ferguson, W. (1993). Epistemic forms and epistemic games: Structures and strategies to guide inquiry. *Educational psychologist*, 28(1), 25-42.
- Cummings, K., Marx, J., Thornton, R., & Kuhl, D. (1999). Evaluating innovation in studio physics. *American journal of physics*, 67(S1), S38-S44.
- d'Andrade, R. G. (1995). *The development of cognitive anthropology*. Cambridge University Press.
- Daniels, H. (Ed.). (2005). *An introduction to vygotsky*. Retrieved from <https://ebookcentral-proquest-com.proxy.lib.utk.edu:2050>

- Derry, J. (2013). Vygotsky - philosophy and education: philosophy and education. Retrieved from <https://ebookcentral-proquest-com.proxy.lib.utk.edu:2050>
- diSessa, A. A. (1996). What do “just plain folk” know about physics? In D. R. Olson & N. Torrance (Eds.), *The handbook of education and human development: New models of learning, teaching, and schooling* (pp. 709–730). Oxford, England: Blackwell.
- diSessa, A. (1993). Towards an epistemology of physics. *Cognition and Instruction*, 10, 105–225.
- diSessa, A. A. (1988). Knowledge in pieces. In G. Forman & P. B. Pufall (Eds.), *Constructivism in the computer age* (pp. 49–70). Hillsdale, NJ: Erlbaum.
- diSessa, A. A, Gillespie, N., & Esterly, J. (2004). Coherence versus fragmentation in the development of the concept of force. *Cognitive Science*, 28, 843–900.
- diSessa, A. A, & Sherin, B. (1998). What changes in conceptual change? *International Journal of Science Education*, 20, 1155–1191.
- diSessa, A. A, & Wagner, J. F. (2005). What coordination has to say about transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multi-disciplinary perspective* (pp. 121–154). Greenwich, CT: Information Age.
- Ding, L. (2014). Verification of causal influences of reasoning skills and epistemology on physics conceptual learning. *Physical Review Special Topics-Physics Education Research*, 10(2), 023101.
- Dowd, J. E. (2012). *Interpreting Assessments of Student Learning in the Introductory Physics Classroom and Laboratory*. Harvard University.
- Duell, O. K., & Schommer-Aikins, M. (2001). Measures of people's beliefs about knowledge and learning. *Educational psychology review*, 13(4), 419-449.
- Elby, A. (1999). Another reason that students learn by rote. *Physics Education Research: A Supplement to the American Journal of Physics*, 67(7), S53-S60.
- Elby, A., & Hammer, D. (2001). On the substance of a sophisticated epistemology. *Science Education*, 85(5), 554-567.
- Greene, J. A., Azevedo, R., & Torney-Purta, J. (2008). Modeling epistemic and ontological cognition: Philosophical perspectives and methodological directions. *Educational Psychologist*, 43(3), 142-160.
- Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. *Handbook of educational psychology*, 77, 15-46.

- Halloun, I. (1998). Views about science and physics achievement. The VASS story. In E. F. Redish & J.S. Rigden (Eds.), *Proceedings of the International Conference on Undergraduate Physics Education* (1996) (pp. 605-613). Washington, DC: American Institute of Physics.
- Hammer, D. (1994). Epistemological beliefs in introductory physics. *Cognition and Instruction*, 12, 151-183.
- Hammer, D. (2000). Student resources for learning introductory physics. *American Journal of Physics*, 68(S1), S52-S59.
- Hammer, D., & Elby, A. (2003). Tapping epistemological resources for learning physics. *The Journal of the Learning Sciences*, 12(1), 53-90.
- Hammer, D., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 89–120). Greenwich, CT: Information Age.
- Hewson, P. W. (1985). Epistemological commitments in the learning of science: Examples from dynamics. *European Journal of Science Education*, 7, 163-172.
- Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of educational research*, 67(1), 88-140.
- Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology*, 25, 378–405.
- Hofer, B. K. (2006). Beliefs about knowledge and knowing: Integrating domain specificity and domain generality: A response to Muis, Bendixen, and Haerle (2006). *Educational Psychology Review*, 18(1), 67–76.
- King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco: Jossey-Bass.
- Kitchener, K. S., & King, P. M. (1981). Reflective judgment: Concepts of justification and their relationship to age and education. *Journal of Applied Developmental Psychology*, 2, 89-116.
- Kuhn, D. (1991). *The skills of argument*. Cambridge, England: Cambridge University Press.
- Langford, P.E. (2005). *Vygotsky's Developmental and Educational Psychology*. Taylor & Francis
- Linn, M. C., & Songer, N. B. (1993). How do students make sense of science? *Merrill-Palmer Quarterly Journal of Developmental Psychology*, 39(1), 47-73.

- Lising, L., & Elby, A. (2005). The impact of epistemology on learning: A case study from introductory physics. *American Journal of Physics*, 73(4), 372-382.
- May, D. B., & Etkina, E. (2002). College physics students' epistemological self-reflection and its relationship to conceptual learning. *American Journal of Physics*, 70(12), 1249-1258.
- McDermott, L. C. (1990). A View from Physics. In M. Gardner, & J. G. Greeno (Eds.), *Toward a Scientific Practice of Science Education* (pp. 3-30). Hillsdale: Lawrence Erlbaum Associates.
- McCloskey, M. (1983a, April). Intuitive physics. *Scientific American*, 122-130.
- McCloskey, M. (1983b). Naive theories of motion. In D. Gentner & A. Stevens (Eds.), *Mental models* (pp. 299-323). Hillsdale, NJ: Erlbaum.
- Muis, K. R., Bendixen, L.D., & Haerle, F. C. (2006). Domain-general and domain-specificity in personal epistemology research: Philosophical and empirical reflections in the development of a theoretical framework. *Educational Psychology Review*, 18, 3-54.
- Otero, V. K. (2004). Cognitive Processes and the Learning of Physics Part 1: The Evolution of Knowledge from a Vygotskian Perspective.
- Otero, V. K. (2004). Cognitive Processes and the Learning of Physics Part 2: Mediated Action.
- Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart and Winston.
- Perry, W. G. (1981). Cognitive and ethical growth: The making of meaning. In A. Chickering (Ed.), *The modern American college* (pp. 76-116). San Francisco: Jossey-Bass.
- Redish, E. F., Steinberg, R. N., & Saul, J. M. (1998). Student expectations in introductory physics. *American Journal of Physics*, 66, 212-224.
- Redish, E. F. (2004). A theoretical framework for physics education research: Modeling student thinking. arXiv preprint physics/0411149.
- Redish, E. F. (2005). Changing student ways of knowing: What should our students learn in a physics class. *Proceedings of World View on Physics Education 2005: Focusing on Change*, New Delhi, 1-13.
- Rieber, R., & SpringerLink. (1999). *The Collected Works of L. S. Vygotsky: Scientific Legacy* (Cognition and language).
- Rieber, R., Robinson, David K, & SpringerLink. (2004). *The Essential Vygotsky*.

- Roth, W. M., & Roychoudhury, A. (1994). Physics students' epistemologies and views about knowing and learning. *Journal of Research in Science Teaching*, 31(1), 5-30.
- Royce, J. R. (1978). Three ways of knowing and the scientific world view. *Methodology and Science*, 11, 146-164.
- Scherr, R. E. (2007). Modeling student thinking: An example from special relativity. *American Journal of Physics*, 75(3), 272-280.
- Scherr, R. E., & Hammer, D. (2009). Student behavior and epistemological framing: Examples from collaborative active-learning activities in physics. *Cognition and Instruction*, 27(2), 147-174.
- Schoenfeld, A. H. (1989). Exploration of students' mathematical beliefs and behavior. *Journal for Research in Mathematics Education*, 20, 338-355.
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of educational psychology*, 82(3), 498.
- Schommer-Aikins, M., Duell, O. K., & Barker, S. (2003). Epistemological beliefs across domains using Biglan's classification of academic disciplines. *Research in Higher Education*, 44, 347-366.
- Schunk, D. H. (2012). *Learning theories an educational perspective sixth edition*. Pearson.
- Shaffer, D. W. (2017). *Quantitative ethnography*. Cathcart Press.
- Shaffer, & Graesser. (2010). Using quantitative model of participation in a community of practice to direct automated mentoring in an ill-formed domain. Paper presented at Intelligent Tutoring Systems (ITS), Pittsburgh, PA.
- Shaffer, Hatfield, Svarovsky, Nash, Nulty, Bagley, Frank, Rupp, & Mislevy, (2009). Epistemic Network Analysis: A prototype for 21st Century assessment of learning. *The International Journal of Learning and Media*. 1(2), pp. 33-53.
- Sherin, B. L. (2000). Meta-representation: An introduction. *The Journal of Mathematical Behavior*, 19(4), 385-398.
- Smith, C. L., Maclin, D., Houghton, C., & Hennessey, M. G. (2000). Sixth-grade students' epistemologies of science: The impact of school science experiences on epistemological development. *Cognition and Instruction*, 18, 349-422.
- Smith, J. I., & Tanner, K. (2010). The problem of revealing how students think: concept inventories and beyond. *CBE-Life Sciences Education*, 9(1), 1-5.

- Stodolsky, S. S., Salk, S., & Glaessner, B. (1991). Student views about learning math and social studies. *American Educational Research Journal*, 28, 89–116.
- Svarovsky, G. N. (2011). Exploring complex engineering learning over time with epistemic network analysis. *Journal of Pre-College Engineering Education Research (J-PEER)*, 1(2), 4.
- Taber, K. S. (2008). Conceptual resources for learning science: Issues of transience and grain- size in cognition and cognitive structure. *International Journal of Science Education*, 30(8), 1027-1053.
- Tannen, D. (1993). What's in a frame? Surface evidence for underlying expectations. In *Framing in discourse*, p.14-56.
- Tuminaro, J. (2004). A cognitive framework for analyzing and describing introductory students' use and understanding of mathematics in physics.
- Tuminaro, J., & Redish, E. F. (2007). Elements of a cognitive model of physics problem solving: Epistemic games. *Physical Review Special Topics-Physics Education Research*, 3(2), 020101.
- Thayer-Bacon, B. J. (2003). Relational"(e)pistemologies". P. Lang.
- Vosniadou, S. (2002). On the nature of naïve physics. In M. Limón & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 61–76). Dordrecht, The Netherlands: Kluwer Academic.
- Vosniadou, S., & Ioannides, C. (1998). From conceptual development to science education: A psychological point of view. *International Journal of Science Education*, 20, 1213–1230.
- Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive psychology*, 24(4), 535-585.
- Vygotsky, L. S. (1997). *Educational psychology*. (V. V. Davydov, Intro.; R. Silverman, Trans.). Boca Raton: St. Lucie Press. (Originally published in Russian, 1926)
- Vygotsky, L., Veer, René van der, & Valsiner, Jaan. (1994). *The Vygotsky reader*. Oxford, UK; Cambridge, Mass., USA: Blackwell.
- Vygotsky, L. S. (1987). *Collected works, Volume1: Problems of general psychology* (R. W Rieber & A. S. Carton, Eds.). New York: Kluwer Academic/Plenum.
- Vygotsky, L. S. (1962). *Thought and language*. Cambridge: M.I.T. Press, Massachusetts Institute of Technology.

- Vygotsky, L. S. (1962). The Development of Scientific Concepts in Childhood. In L. Vygotsky & E. Hanfmann, G. Vakar (Eds.), *Studies in communication. Thought and language* (pp. 82-118). Cambridge, MA, US: MIT Press. <http://dx.doi.org.proxy.lib.utk.edu/90/10.1037/11193-006>
- Wagner, J. F. (2006). Transfer in pieces. *Cognition and Instruction*, 24, 1–71.
- Wagner, J. F. (Chair). (2005, April). On the nature of students' knowledge: Contrasting epistemologies in science and mathematics education research. Symposium organized by J. Wagner including D. Clark, A. diSessa, J. Mestre, S. Vosniadou, and J. Wagner for the American Educational Research Association Annual Conference 2005, Montreal, Quebec, Canada.

Chapter 3: Methodology

- Adams, W. K., & Wieman, C. E. (2011). Development and validation of instruments to measure learning of expert- like thinking. *International Journal of Science Education*, 33(9), 1289-1312.
- Ahmed, W. (July 10, 2015). Using twitter as a data source: An overview of current social media research tools. [Web log entry] Retrieved from <http://blogs.lse.ac.uk/impactofsocialsciences/2015/07/10/social-media-research-tools-overview/>
- Bengtsson, M. (2016). How to plan and perform a qualitative study using content analysis. *Nursing Plus Open*, 2, 8-14.
- Berelson, B. L. (1952). *Content analysis in communications research*. New York: Free press.
- Borra, E., & Rieder. (2014) "Programmed method: developing a toolset for capturing and analyzing tweets," *Aslib Journal of Information Management*, Vol. 66 (3), pp.262 - 278. <http://dx.doi.org/10.1108/AJIM-09-2013-0094>
- Brooker, P., Barnett, J., & Cribbin, T. (2016). Doing social media analytics. *Big Data & Society*, 3(2).
- Burghardt, M. (2015) Introduction to tools and methods for the analysis of twitter data. *10plus1: Living Linguistics, Media Linguistics*, 1.
- Burns N. & Grove S.K. (2005) *The Practice of Nursing Research: Conduct, Critique & Utilization*. Elsevier Saunders, St Louis.
- Caracelli, V. J., & Greene, J. C. (1993). Data analysis strategies for mixed-method evaluation designs. *Educational evaluation and policy analysis*, 15(2), 195-207.
- Chen, L., & Chen, T. L. (2012). Use of Twitter for formative evaluation: Reflections on trainer and trainees' experiences. *British Journal of Educational Technology*, 43(2), E49-E52.

- Chew, C., & Eysenbach, G. (2010). Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PloS one*, 5(11), e14118.
- Collins, K. M., Onwuegbuzie, A. J., & Sutton, I. L. (2006). A model incorporating the rationale and purpose for conducting mixed methods research in special education and beyond. *Learning disabilities: a contemporary journal*, 4(1), 67-100.
- Creswell, J. W. (2015). *A concise introduction to mixed methods research*. Sage Publications.
- Creswell, J. W., & Clark, V. L. P. (2007). *Designing and conducting mixed methods research*.
- Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. *Handbook of mixed methods in social and behavioral research*, 209-240.
- Denzin, N. K., & Giardina, M. D. (Eds.). (2016). *Qualitative inquiry and global crises*. Routledge.
- Downe-Wambolt, B. (1992). Content analysis: method, applications and issues. *Health Care for Women International*, 13, 313-321.
- Dunlap, J. C., & Lowenthal, P. R. (2009). Tweeting the night away: Using Twitter to enhance social presence. *Journal of Information Systems Education*, 20(2), 129.
- Duschl, R., Schweingruber, H., & Shouse, A., (Eds.) (2007). *Taking science to school. Learning and teaching science in Grades K-8*. Washington, DC: National Academies Press.
- Elby, A., Frederiksen, J., Schwarz, C., and White, B. 2001, "The Epistemological Beliefs Assessment for Physical Science," www.flagguide.org
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of advanced nursing*, 62(1), 107-115.
- Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of educational research*, 82(3), 300-329.
- Gesink, D., Filsinger, B., Mihic, A., Norwood, T. A., Racey, C. S., Perez, D., ... & Vernich, L. (2016). Cancer screening barriers and facilitators for under and never screened populations: A mixed methods study. *Cancer Epidemiology*, 45, 126-134.
- Graneheim, U.H., & Lundman, B. (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, 24, 105–112.

- Gruzd, A. (2016). Netlytic: Software for Automated Text and Social Network Analysis. Available at <http://Netlytic.org>
- Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research*, 15(9), 1277-1288.
- Junco, R., Heiberger, G., & Loken, E. (2011). The effect of Twitter on college student engagement and grades. *Journal of Computer Assisted Learning*, 27, 119–132.
- Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of mixed methods research*, 1(2), 112-133.
- Krippendorff, K. (2004). *Content analysis: an introduction to its methodology*. Thousand oaks, California: Sage Publications Inc.
- Leech, N. L., & Onwuegbuzie, A. J. (2007). An array of qualitative analysis tools: A call for data analysis triangulation. *School Psychology Quarterly*, 22, 557–584. doi:10.1037/1045–3830.22.4.557.
- Lincoln, Y. S. & Guba, E. G. (1985). *Naturalistic Inquiry*. Sage Publications, Newbury Park, CA.
- Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative research: A guide to design and implementation*. John Wiley & Sons.
- Minstrell, J., Anderson, R., & Li, M. (2011). Building on learner thinking: A framework for assessment in instruction. Commissioned paper for the Committee on Highly Successful STEM Schools or Programs for K-12 STEM Education.
- Osborne, J. F., Duschl, R., & Fairbrother, R. (2002). *Breaking the mould: Teaching science for public understanding*. Nuffield Foundation.
- Patton, M. Q. (2005). *Qualitative research*. John Wiley & Sons, Ltd.
- Paul, J. E., & Iannitti, N. (2012). On beyond clickers: Twitter as a classroom response system. *Journal of Health Administration Education*, 29(4), 319-328.
- Perkins, K., Adams, W., Finkelstein, N., LeMaster, R., Reid, S., Dubson, M., ... & Wieman, C. (2004). The Physics Education Technology Project: A New Suite of Physics Simulations. In Poster Presented at AAPT Summer Meeting.
- Polit D.F. & Beck C.T. (2004) *Nursing Research. Principles and Methods*. Lippincott Williams & Wilkins, Philadelphia, PA.
- Potter, W. J., & Levine- Donnerstein, D. (1999). Rethinking validity and reliability in content analysis.
- Powell, T. C. (2001) Competitive advantage: logical and philosophical considerations, *Strategic Management Journal*, 22(9), pp. 875–88.

- Rankin, M. (2012). The Twitter experiment at UT Dallas. Web. < <http://www.utdallas.edu/~mrankin/usweb/twitterconclusions.htm>.
- Redish, E. F., Saul, J. M., & Steinberg, R. N. (1998). Student expectations in introductory physics. *American Journal of Physics*, 66(3), 212-224.
- Sample, M. (2010). A framework for teaching with Twitter. *Chronicle of Higher Education*.
- Saldaña, J. (2009). *The coding manual for qualitative researchers*.
- Sandoval, W., & Reiser, B. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education*, 88, 345-372.
- Schreier, M. (2012). *Qualitative content analysis in practice*. Sage Publications.
- Shulman, S. (2011, June). DiscoverText: Software training to unlock the power of text. In *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times* (pp. 373-373). ACM.
- Sullivan, G. M. (2011). A Primer on the Validity of Assessment Instruments. *Journal of Graduate Medical Education*, 3(2), 119–120. <http://doi.org/10.4300/JGME-D-11-00075.1>
- Teddlie, C., & Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in the social and behavioral sciences. *Handbook of mixed methods in social & behavioral research*, 3-50.
- Teddlie, C., & Tashakkori, A. (2006). A general typology of research designs featuring mixed methods. *Research in the Schools*, 13(1), 12-28.
- Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Sage.
- Terrell, S. R. (2012). Mixed-methods research methodologies. *The qualitative report*, 17(1), 254-280.
- Tiernan, P. (2014). A study of the use of Twitter by students for lecture engagement and discussion. *Education and Information Technologies*, 19(4), 673-690.
- Trochim, W. M. (2006). Qualitative validity. *Research methods knowledge base*, 1-3.
- Trochim, W., Donnelly, J. P., & Arora, K. (2015). *Research methods: The essential knowledge base*. Nelson Education.
- Young, J. R. (2010). Teaching with twitter: Not for the faint of heart. *Education Digest: Essential Readings Condensed for Quick Review*, 75(7), 9-12.

Chapter 4: Findings and Results

- Ernest, P. (2003). *Constructing mathematical knowledge: Epistemology and mathematics education*. Routledge.
- McMillan, James. (2013). *SAGE Handbook of Research on Classroom Assessment*. Thousand Oaks: SAGE Publications.
- Means, B., Chen, E., DeBarger, A., & Padilla, C. (2011). *Teachers' Ability to Use Data to Inform Instruction: Challenges and Supports*. Office of Planning, Evaluation and Policy Development, US Department of Education.
- Nitko, A. J., & Brookhart, S. M. (2007). *Educational Assessment of Students* (5th ed.). Columbus, OH: Pearson Prentice Hall.
- Waggoner, J., & Carroll, J. B. (2014). Concurrent validity of standards-based assessments of teacher candidate readiness for licensure. *SAGE Open*, 4(4), 2158244014560545.

Chapter 5: Discussion and Conclusions

- Andrist, S., Collier, W., Gleicher, M., Mutlu, B., & Shaffer, D.W. (2015). Look together: Analyzing gaze coordination with epistemic network analysis. *Frontiers in Psychology*, 6(1016).
- Beichner, R. (2008). *The SCALE-UP Project: a student-centered active learning environment for undergraduate programs*. An invited white paper for the National Academy of Sciences.
- Belenky, M. R., Clinchy, B. M., Goldberger, N. R., & Tarule, J. M. (1986). *Women's ways of knowing: The development of self, voice and mind*. New York: Basic Books.
- Belenky, M., Clinchy, B., Goldberger, N., & Tarule, J. (1996). *Knowledge, difference, and power: Essays inspired by women's ways of knowing*. New York: Basic.
- Brewe, E., Bartley, J. E., Riedel, M. C., Sawtelle, V., Salo, T., Boevig, E. R., ... & Laird, R. W. (2018). Toward a neurobiological basis for understanding learning in University Modeling Instruction physics courses. *Frontiers in ICT*, 5, 10.
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American journal of Physics*, 66(1), 64-74.
- Halpern, D. F., Aronson, J., Reimer, N., Simpkins, S., Star, J. R., & Wentzel, K. (2007). *Encouraging girls in math and science* (NCER 2007-2003). Washington, DC: U.S. Department of Education, National Center for Education Research.

- Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, 269, 41–45.
- Ivanjek, L., Susac, A., Planinic, M., Andrasevic, A., & Milin-Sipus, Z. (2016). Student reasoning about graphs in different contexts. *Physical Review Physics Education Research*, 12(1), 010106.
- Kimura, D. (2002). Sex hormones influence human cognitive pattern. *Neuroendocrinology Letters*, 23(Suppl. 4), 67–77.
- Kozhevnikov, M., Hegarty, M., & Mayer, R. E. (2002). Revising the visualizer-verbalizer dimension: Evidence for two types of visualizers. *Cognition and Instruction*, 20, pp. 47–78.
- Kozhevnikov, M., and Thornton, R. (2006). Real-Time data display, spatial visualization ability, and learning force and motion concepts. *Journal of Science Education and Technology*, 15(1), 111-132.
- McDermott, L. C., Rosenquist, M. L., & Van Zee, E. H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics*, 55(6), 503-513.
- Minstrell, J., Anderson, R., & Li, M. (2011). Building on learner thinking: A framework for assessment in instruction. Commissioned paper for the Committee on Highly Successful STEM Schools or Programs for K-12 STEM Education.
- Philipp, S. B., Tretter, T. R., & Rich, C. V. (2016). Undergraduate teaching assistant impact on student academic achievement. *Electronic Journal of Science Education*, 20(2).
- Ruis, A.R., Rosser, A.A., Quandt-Walle, C., Nathwani, J.N., Shaffer, D.W., & Pugh, C.M. (2017). The hands and head of a surgeon: Modeling operative competency with multimodal epistemic network analysis. *The American Journal of Surgery*. Advance online publication.
- Rus, V., Gautam, D., Swiecki, Z., Shaffer, D. W., & Graesser, A. (2016). Assessing Student-Generated Design Justifications in Virtual Engineering Internships. In EDM (pp. 496-501).
- Sellami, N., Shaked, S., Laski, F. A., Eagan, K. M., & Sanders, E. R. (2017). Implementation of a Learning Assistant Program Improves Student Performance on Higher-Order Assessments. *CBE Life Sciences Education*, 16(4), ar62. <http://doi.org/10.1187/cbe.16-12-0341>
- Severiens, S.E., & Geert, T.M. (1994). Gender differences in learning styles: a narrative review and quantitative meta-analysis. *Higher Education*, 27, pp. 487-501.

- Sharma, M. D., Johnston, I. D., Johnston, H., Varvell, K., Robertson, G., Hopkins, A., Stewart, C., Cooper, I., & Thornton, R. (2010). Use of interactive lecture demonstrations: A ten-year study. *Physical Review Special Topics-Physics Education Research*, 6(2), 020119.
- Stadler, H., Duit, R. and Benke, G. (2000) Do boys and girls understand physics differently? *Physics Education*, 35, 417–422.
- Stang, J. B., & Roll, I. (2014). Interactions between teaching assistants and students boost engagement in physics labs. *Physical Review Special Topics-Physics Education Research*, 10(2), 020117.
- Thomsen, T., Hugdahl, K., Ersland, L., Barndon, R., Lundervold, A., Smievoll, A. I., Roscher, B. E., & Sundberg, H. (2000). Functional magnetic resonance imaging (fMRI) study of sex differences in a mental rotation task. *Medical Science Monitor*, 6(6), 1186-1196.
- Unterrainer, J., Wranek, U., Staffen, W., Gruber, T., & Ladurner, G. (2000). Lateralized cognitive visuospatial processing: Is it primarily gender-related or due to quality of performance? *Neuropsychobiology*, 41(2), 95-101.
- Was, C., Sansosti, F., & Morris, B. (Eds.). (2016). *Eye-Tracking Technology Applications in Educational Research*. IGI Global.
- Weiss, E. M., Kemmler, G., Deisenhammer, E. A., Fleischhacker, W. W., & Delazer, M. (2003). Sex differences in cognitive functions. *Personality and individual differences*, 35(4), 863-875.
- Wilcox, M., Feldman, G., Von Korff, J., Klinger, N., Sengul, O., & Chini, J. (2016). Characterizing studio physics instruction across instructors and institutions. In *Proceedings of the 2016 Physics Education Research Conference*, Sacramento, CA (p. 392).

APPENDIX

Appendix 1A: Experiment 1 Lab

The resistance of your body: Fatal electric shock occurs when a sufficiently large electric current flows through the body. A fraction of such a current flows through the heart and may disrupt the cardiac cycle. Typical effects are listed in the table below.

Shocking current:	Effect:
< 1mA	no observable effect
~1mA - ~10 mA	tingling sensation
~10 mA - ~100 mA	muscular paralysis ("can't let go")
~100 mA	ventricular fibrillation
~1A - ~ 10 A	thermal damage to tissue

Paradoxically, brief currents of > 1 A may be less dangerous than lower currents. Instead of putting the heart into ventricular fibrillation, these currents clamp the whole heart muscle at the same time. When the current is turned off, a normal heart beat may resume on its own accord. Indeed, currents of about 1 A are used clinically to defibrillate the heart.

Experiment 1

Use the digital multimeter to measure the resistance of your body. Switch the meter on the $20\text{ M}\Omega$ scale. Make sure the leads are plugged into the Ω and COM connectors on the lower right of the meter.

Press the thumb of one of your hands against the black and the thumb of the other hand against the red lead. Record the values for each member of your group with dry and with wet thumbs.

- Predict the current that would flow through your body if you held one terminal of a 1.5 V battery in one of your hands and the other terminal in the other hand.
- Predict the current that would flow through your body if you held one terminal of a 110 V power source in one of your hands and the other terminal in the other hand.

The salty fluids within the human body are electrical conductors. The internal resistance of an arm (from hand to shoulder) is less than $100\ \Omega$. If there is a voltage across this internal resistance, a current will flow and heat will be generated. If the current is large or the connection time is long enough, this heat will cause burns and destroy tissue. Fortunately, the resistance of dry skin is high. Using a typical contact area, the skin acts like an approximately (10 - 100) $\text{k}\Omega$ resistor in series with the internal



resistance of the body. At voltages below about 50 V the **dry** skin provides safe current limiting protection.

Be extremely careful not to have electrical contact with a voltage source if you have wet or sweaty skin.

Appendix 1B: Experiment 2 Lab

Measuring capacitance

Assume you connect two identical metal plates of area A , separated by a non-conducting material which has a thickness d to a battery and a switch, as shown. When the switch is open, there is no excess charge on either plate. Discuss what happens when the switch is then closed.

Experiment 2

Construct a parallel plate capacitor out of two rectangular pieces of metal foil. The sheets have small "handles". Attach the leads of the multimeter to those handles and make sure the leads are plugged into the $+$ and $-$ connectors on the lower left of the meter. Switch the multimeter to the 20 nF scale.

Slip the two foil sheets between the pages of a heavy textbook and separate them by 3 pages of the book. Make sure the foil sheets do not touch each other and "short out". The areas of the sheets should overlap and the "handles" should stick out on opposite sides. Weight the book down with another heavy textbook.

- Measure the capacitance of your "parallel plate capacitor" using your multimeter. Record your measured value.
- Investigate how the capacitance depends on the separation between the two foil sheets. Take data for at least five different separation distances (in units of number of pages). Make sure everything except the separation of the two foils stays the same. Record your data in a table in your log. Record the dimensions of the foil and calculate its area.
- Use Excel to produce a plot of capacitance C in units of nF versus d in units of number of pages and versus $1/d$. Paste the graphs into your log. Is either of these graphs linear?

number of pages:	Capacitance (nF)
3	
6	
9	
12	
15	

- Investigate how the capacitance depends on the area of the conducting plates. Use a separation distance of 3 pages and change the overlap area of the plates to approximately $1/2$ and $1/4$ of the total area. Just move the foils, do **NOT** fold them. Record your data in a table in your log.



Appendix 1C: Experiment 3 Lab

Ohm's law? Ohm discovered that when the voltage (potential difference) across a conductor changes, the current flowing through the conductor changes. He expressed this as $I = V/R$ or $V = IR$. As the voltage increases, so does the current. For many conductors, R is approximately constant. These materials are called ohmic. If the voltage across an ohmic resistor is increased, a plot of voltage versus current shows a straight line. The slope of the line is equal to R . For non-ohmic materials, R is not constant and a plot of voltage versus current will not show a straight line.

Experiment 3

Use the Pasco RLC board to investigate the relationship between current and voltage in ohmic and non-ohmic materials.

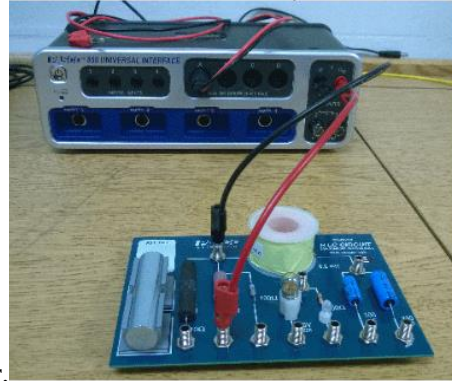
- Make sure the Pasco 850 interface is turned on. Open the Capstone program. The icon for this program is on the desktop.
- In the Capstone program, click the Hardware Setup button on the left and add a Science Workshop Analog Sensor. Choose the Voltage Sensor.
- Plug the Voltage Sensor into Analog Channel A, but move the leads to the side. You will only use this sensor later.
- In the Hardware Setup window click on the yellow circle over the two output jacks on the image of the PASCO 850 interface.



Select the Output Voltage - Current Sensor.

- Plug a black patch cord into the left output and a red patch cord into the right output.
- Click Hardware Setup again to close this window. In the Control Panel at the bottom set the sample rate to 1000 Hz = 1 kHz.
- Click on the Signal Generator icon in the tool panel, (4 icons below Hardware Setup), and choose the 850 Output 1.
- Choose the Triangle Wave, set the amplitude to 3 V, the frequency to 1 Hz. Click AUTO to turn it on automatically.
- Close the signal generator window and click on "Two Small, One Large Display".
- For the vertical axis of one small graph choose Output Voltage (V). Make sure you do NOT choose Voltage (Ch A). For the vertical axis of the other small graph choose Output Current (A). For both small graphs choose time for the horizontal axis.
- For the large graph choose Output Voltage (V) for the vertical and Output Current (A) for the horizontal axis.
- Click the Recording Conditions icon in the control panel.
Choose Start Condition, Measurement Based, **Output Voltage (V)** is above 0.2 V. Make sure you do NOT choose Voltage (Ch A).
Choose Stop Condition, Time Based, 3 s.

- Connect the $33\ \Omega$ resistor to the output of the interface as shown below. Use the black and a red



patch cords, not the leads from the voltage sensor.

-
-
- Click the Record button.
 - You will record data for 3 seconds.
 - Click the "autoscale" button in the graph displays to rescale the graphs.
 - In V versus A graph click Fit, Linear Fit. The slope of the large V versus A graph will be displayed.
- Give your large graph of voltage versus current a title and
- copy it into your Word d document. Repeat the measurement with the $100\ \Omega$ resistor connected to the output of the interface.
- Answer the following questions:
 - Are the $33\ \Omega$ and $100\ \Omega$ resistor ohmic resistors? Why or why not? Base your conclusions on your measurements.
 - Do the resistances found from the slope of the voltage versus current graphs agree with the nominal resistances written on the RLC board? What is the percentage difference?
- Now erase all your data and connect the light bulb on the RLC board to the output of the interface and take data for 3 s.

Repeat the measurement a few times, and alternately watch the bulb and the trace on your large graph.

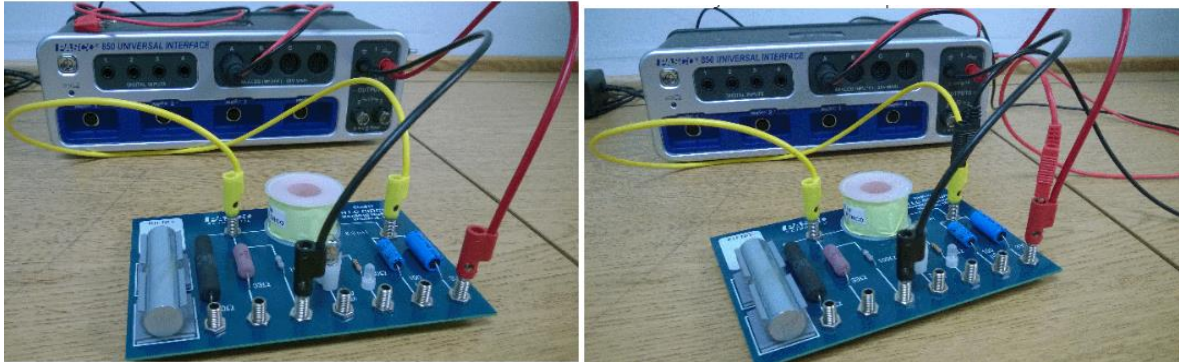
- Give your large graph of voltage versus current a title and copy it into your Word document.
- Answer the following questions
 - Is the bulb an ohmic resistor? Why or why not? Base your conclusions on your measurements. Describe your observations.
 - Can you determine the resistance of the bulb?
- Now erase all your data and connect the LED on the RLC board to the output of the interface and take data for 3 s.
- - Repeat the measurement a few times, and alternately watch the LED and the trace on the graph 3.
- Describe your observations.
- Change the signal generator amplitude to 5 V and repeat.
- Describe your observations.
- Give your large graph of voltage versus current a title and copy it into your Word document.
- Is the LED an ohmic resistor? Why or why not? Base your conclusions on your measurements.

Appendix 1D: Experiment 4 Lab

RC circuits

Find out how the voltage across a capacitor varies as it charges and discharges.

Experiment 4 Connect the 300 μF capacitor and the 100 Ω resistor in series to the output of the interface as shown below. Note: You have to short out the coil.



- In the Signal Generator window choose the square wave.
- Choose an amplitude of 2 V and a frequency of 2 Hz.
Under Offsets and Limits adjust the voltage offset to be 2 volts.
- Connect the voltage sensor across the capacitor as shown below.
- Change the axes of your large graph. Choose Voltage (Ch A) for the vertical axis and time for the horizontal axis.
- Click the Record button. You will record data for 3 s. The power supply will turn on for 1/4 s, then turn off for 1/4 s then turn on again for 1/4 s and so on. During the time the power supply is off the RC circuit is shorted out. The capacitor will charge when the power supply is on and discharge when it is off.
- Click the "autoscale" button in your graphs to rescale the display. Magnify a region of the plot of voltage versus time that shows the voltage rising from zero volts to the maximum voltage.
 - Use the smart cursor to find the time t_1 when the voltage begins to rise and the time t_2 when the voltage has reached 1/2 of its maximum value.
Find $t_{1/2} = t_2 - t_1$.
 - Use $t_{1/2} = \tau \ln 2 = 0.693 \tau$ to find the time constant τ .
- Answer the following questions:
 - What value did you find for τ ? Does this value agree with the nominal value $\tau = RC$?
 - Note: The stated values of the capacitance may vary by as much as 20% from the actual value.
 - Check your value of $t_{1/2}$ obtained from the voltage versus time plot by also obtaining it from the output current versus time plot. How long does it take for the current to decrease to 1/2 its maximum value? do you get the same value for $t_{1/2}$?

Appendix 2A: Study Information

Using Twitter and Network Analysis to Assess the Development of Complex Thinking in Physics: An Exploratory Mixed Methods Study

PURPOSE

The purpose of this research study is to learn how you make sense of and think about inquiry tasks, using Twitter to collect data on your thoughts, as you complete inquiry activities.

PROCEDURES

1. The research activities will occur during class time of Studios 2, 3 and 4. Detailed Study Information and Instructions will be available on Canvas for your reference.
2. Before Studio 2, you will receive ‘Instructions’ to set up your individual Twitter Account, using your NetID.
3. During Studios 2, 3 and 4, when the class divides into groups, you will work with members who are also participating in the research.
4. During Studio 2, you will complete a questionnaire called the ‘Epistemological Beliefs about Physical Sciences Survey’. It will take 15 to 20 minutes and is used to determine your prior beliefs and expectations about physics.
5. During Studio 2, 3, and 4, you will use Twitter in two ways.
 - a. You will Tweet your responses to specific prompts and questions related to the concepts, activities and tasks;
 - b. You will use Twitter instead of spoken words, to communicate your thoughts and questions about strategy and difficulties with your group members and the TAs during inquiry activities.
6. This research also includes the use of your course materials and information such as Clicker quiz response data, homework quiz response data, and your lab report scores of Studios 3 and 4; and Test 1 scores.

RISKS

Most studies involve some risk to confidentiality and it is possible that someone could find out you were in this study or see your study information, but the researcher believes this risk is unlikely to occur because of the procedures used to protect your information.

BENEFITS

Although we do not anticipate that you will directly benefit from participating in this research, you may find that using Twitter to reflect on and express your thoughts about the coursework helps you with the course. This study may provide valuable information for making physics accessible to all students.

CONFIDENTIALITY

All data collected will be kept confidential. Data will be stripped of all identifiers, by replacing your usernames and twitter names with pseudonyms, per a coding key, once all data has been linked. The coding key and all digital data will be stored securely on a pass-word protected computer, and will be

made available only to persons conducting the study. The coding key will be retained on the computer for three years from the date of its creation. No reference will be made in oral or written reports published in journals, which could link you to the study. Digital data, stripped of identifiers, and the coding key will be backed up to the cloud storage system, One Drive for Business of the University of Tennessee, and retained for a period of three years. After three years, the data will be deleted as per University guidelines.

COMPENSATION

All who volunteer to participate in this study will earn extra credit in the course. Those who choose not to participate will still have extra credit opportunities available to them in the form of additional lab or homework assignments. There is no penalty for withdrawing from the study at any time. Full credit will be given for participating in the study during studio sessions 3 and 4; and 50% for participating in either studio 3 or 4.

CONTACT INFORMATION

If you have questions at any time about the study or the procedures, or you experience adverse effects because of participating in this study, you may contact the researcher, Deepa Deshpande at ddeshpan@vols.utk.edu, or by telephone at 865-974-8145. If you have questions about your rights as a participant, contact the University of Tennessee, IRB Compliance Officer and update our contact information to 865-974-7697 and utkirb@utk.edu.

PARTICIPATION

Participants must be age 18 or older to participate. Your participation in this study is voluntary. Your decision to participate or not participate in this study, or discontinue/stop your participation will not affect your relationship with the course instructors or your course grade in any way. If you withdraw from the study your data that has already been collected will be used for the research.

Appendix 2B: Informed Consent

INFORMED CONSENT

I have received, read and fully understand the Study Information. I acknowledge that my participation in the study is voluntary, that I am 18 years of age, and I am aware that I may choose to terminate my participation in the study at any time and for any reason. I agree to participate in the research study “Using Twitter and Network Analysis to Assess the Development of Complex Thinking in Physics: An Exploratory Mixed Methods Study”.

Participant's Name (printed) _____

Participant's Signature _____ Date _____

Appendix 3A: EBAPS Survey

Part 1

DIRECTIONS: For each of the following items, please read the statement, and indicate (on the scantron answer sheet) the answer that describes how strongly you agree or disagree.

A: Strongly disagree B: Somewhat disagree C: Neutral D: Somewhat agree E: Strongly agree

1. Tamara just read something in her science textbook that seems to disagree with her own experiences. But to learn science well, Tamara shouldn't think about her own experiences; she should just focus on what the book says.
2. When it comes to understanding physics or chemistry, remembering facts isn't very important.
3. Obviously, computer simulations can predict the behavior of physical objects like comets. But simulations can also help scientists estimate things involving the behavior of *people*, such as how many people will buy new television sets next year.
4. When it comes to science, most students either learn things quickly, or not at all.
5. If someone is having trouble in physics or chemistry class, studying in a better way can make a big difference.
6. When it comes to controversial topics such as which foods cause cancer, there's no way for scientists to evaluate which scientific studies are the best. Everything's up in the air!
7. A teacher once said, "I don't *really* understand something until I teach it." But actually, teaching doesn't help a teacher understand the material better; it just reminds her of how much she already knows.
8. Scientists should spend almost all their time gathering information. Worrying about theories can't really help us understand anything.
9. Someone who doesn't have high natural ability can still learn the material well even in a hard chemistry or physics class.
10. Often, a scientific principle or theory just doesn't make sense. In those cases, you have to accept it and move on, because not everything in science is supposed to make sense.
11. When handing in a physics or chemistry test, you can generally have a sense of how well you did even before talking about it with other students.

A: Strongly disagree B: Somewhat disagree C: Neutral D: Somewhat agree E: Strongly agree

12. When learning science, people can understand the material better if they relate it to their own ideas.
13. If physics and chemistry teachers gave *really clear* lectures, with plenty of real-life examples and sample problems, then most good students could learn those subjects without doing lots of sample questions and practice problems on their own.
14. Understanding science is really important for people who design rockets, but not important for politicians.
15. When solving problems, the key thing is knowing the methods for addressing each particular type of question. Understanding the “big ideas” might be helpful for specially-written problems, but not for most regular problems.
16. Given enough time, almost everybody could learn to think more scientifically, if they really wanted to.
17. To understand chemistry and physics, the formulas (equations) are really the main thing; the other material is mostly to help you decide which equations to use in which situations.

Part 2

DIRECTIONS: *Multiple choice. On the answer sheet, fill in the answer that best fits your view.*

18. If someone is trying to learn physics, is the following a good kind of question to think about?

Two students want to break a rope. Is it better for them to (1) grab opposite ends of the rope and pull (like in tug-of-war), or (2) tie one end of the rope to a wall and both pull on the other end together?

- (a) *Yes, definitely.* It’s one of the best kinds of questions to study.
- (b) *Yes, to some extent.* But other kinds of questions are equally good.
- (c) *Yes, a little.* This kind of question is helpful, but other kinds of questions are more helpful.
- (d) *Not really.* This kind of question isn’t that great for learning the main ideas.
- (e) *No, definitely not.* This kind of question isn’t helpful at all.

19. Scientists are having trouble predicting and explaining the behavior of thunder storms. This could be because thunder storms behave according to a very complicated or hard-to-apply set of rules. Or, that could be because some thunder storms don't behave consistently according to *any* set of rules, no matter how complicated and complete that set of rules is.

In general, why do scientists sometimes have trouble explaining things? Please read all options before choosing one.

- (a) Although things behave in accordance with rules, those rules are often complicated, hard to apply, or not fully known.
- (b) Some things just don't behave according to a consistent set of rules.
- (c) Usually it's because the rules are complicated, hard to apply, or unknown; but sometimes it's because the thing doesn't follow rules.
- (d) About half the time, it's because the rules are complicated, hard to apply, or unknown; and half the time, it's because the thing doesn't follow rules.
- (e) Usually it's because the thing doesn't follow rules; but sometimes it's because the rules are complicated, hard to apply, or unknown.

20. In physics and chemistry, how do the most important formulas relate to the most important concepts? Please read all choices before picking one.

- (a) The major formulas summarize the main concepts; they're not really separate from the concepts. In addition, those formulas are helpful for solving problems.
- (b) The major formulas are kind of "separate" from the main concepts, since concepts are *ideas*, not equations. Formulas are better characterized as problem-solving tools, without much conceptual meaning.
- (c) Mostly (a), but a little (b).
- (d) About half (a) and half (b).
- (e) Mostly (b), but a little (a).

21. To be successful at *most things in life*...

- (a) Hard work is much more important than inborn natural ability.
- (b) Hard work is a little more important than natural ability.
- (c) Natural ability and hard work are equally important.
- (d) Natural ability is a little more important than hard work.
- (e) Natural ability is much more important than hard work.

22. To be successful at *science*...

- (a) Hard work is much more important than inborn natural ability.
- (b) Hard work is a little more important than natural ability.
- (c) Natural ability and hard work are equally important.
- (d) Natural ability is a little more important than hard work.
- (e) Natural ability is much more important than hard work.

23. Of the following test formats, which is best for measuring how well students understand the material in physics and chemistry? Please read each choice before picking one.

- (a) A large collection of short-answer or multiple-choice questions, each of which covers one specific fact or concept.
- (b) A small number of longer questions and problems, each of which covers several facts and concepts.
- (c) Compromise between (a) and (b), but leaning more towards (a).
- (d) Compromise between (a) and (b), favoring both equally.
- (e) Compromise between (a) and (b), but leaning more towards (b).

Part 3

DIRECTIONS: *In each of the following items, you will read a short discussion between two students who disagree about some issue. Then you'll indicate whether you agree with one student or the other*

24.

Brandon: A good science textbook should show how the material in one chapter relates to the material in other chapters. It shouldn't treat each topic as a separate "unit," because they're not really separate.

Jamal: But most of the time, each chapter is about a different topic, and those different topics don't always have much to do with each other. The textbook should keep everything separate, instead of blending it all together.

With whom do you agree? Read all the choices before circling one.

- (a) I agree almost entirely with Brandon.
- (b) Although I agree more with Brandon, I think Jamal makes some good points.
- (c) I agree (or disagree) equally with Jamal and Brandon.
- (d) Although I agree more with Jamal, I think Brandon makes some good points.
- (e) I agree almost entirely with Jamal.

25.

Anna: I just read about Kay Kinoshita, the physicist. She sounds naturally brilliant.

Emily: Maybe she is. But when it comes to being good at science, hard work is more important than "natural ability." I bet Dr. Kinoshita does well because she has worked really hard.

Anna: Well, maybe she did. But let's face it, some people are just smarter at science than other people. Without natural ability, hard work won't get you anywhere in science!

- (a) I agree almost entirely with Anna.
- (b) Although I agree more with Anna, I think Emily makes some good points.
- (c) I agree (or disagree) equally with Anna and Emily.
- (d) Although I agree more with Emily, I think Anna makes some good points.
- (e) I agree almost entirely with Emily.

26.

Justin: When I'm learning science concepts for a test, I like to put things in my own words, so that they make sense to me.

Dave: But putting things in your own words doesn't help you learn. The textbook was written by people who know science really well. You should learn things the way the textbook presents them.

- (a) I agree almost entirely with Justin.
- (b) Although I agree more with Justin, I think Dave makes some good points.
- (c) I agree (or disagree) equally with Justin and Dave.
- (d) Although I agree more with Dave, I think Justin makes some good points.
- (e) I agree almost entirely with Dave.

27.

Julia: I like the way science explains things I see in the real world.

Carla: I know that's what we're "supposed" to think, and it's true for many things. But let's face it, the science that explains things we do in lab at school can't really explain earthquakes, for instance. Scientific laws work well in some situations but not in most situations.

Julia: I still think science applies to almost all real-world experiences. If we can't figure out how, it's because the stuff is very complicated, or because we don't know enough science yet.

- (a) I agree almost entirely with Julia.
- (b) I agree more with Julia, but I think Carla makes some good points.
- (c) I agree (or disagree) equally with Carla and Julia.
- (d) I agree more with Carla, but I think Julia makes some good points.
- (e) I agree almost entirely with Carla.

28.

Leticia: Some scientists think the dinosaurs died out because of volcanic eruptions, and others think they died out because an asteroid hit the Earth. Why can't the scientists agree?

Nisha: Maybe the evidence supports both theories. There's often more than one way to interpret the facts. So, we have to figure out what the facts mean.

Leticia: I'm not so sure. In stuff like personal relationships or poetry, things can be ambiguous. But in science, the facts speak for themselves.

- (a) I agree almost entirely with Leticia.
- (b) I agree more with Leticia, but I think Nisha makes some good points.
- (c) I agree (or disagree) equally with Nisha and Leticia.
- (d) I agree more with Nisha, but I think Leticia makes some good points.
- (e) I agree almost entirely with Nisha.

29.

Jose: In my opinion, science is a little like fashion; something that's "in" one year can be "out" the next. Scientists regularly change their theories back and forth.

Miguel: I have a different opinion. Once experiments have been done and a theory has been made to explain those experiments, the matter is pretty much settled. There's little room for argument.

- (a) I agree almost entirely with Jose.
- (b) Although I agree more with Jose, I think Miguel makes some good points.
- (c) I agree (or disagree) equally with Miguel and Jose.
- (d) Although I agree more with Miguel, I think Jose makes some good points.
- (e) I agree almost entirely with Miguel.

30.

Jessica and Mia are working on a homework assignment together...

Jessica: O.K., we just got problem #1. I think we should go on to problem #2.

Mia: No, wait. I think we should try to figure out why the thing takes so long to reach the ground.

Jessica: Mia, we know it's the right answer from the back of the book, so what are you worried about? If we didn't understand it, we wouldn't have gotten the right answer.

Mia: No, I think it's possible to get the right answer without really understanding what it means.

- (a) I agree almost entirely with Jessica.
- (b) I agree more with Jessica, but I think Mia makes some good points.
- (c) I agree (or disagree) equally with Mia and Jessica.
- (d) I agree more with Mia, but I think Jessica makes some good points.
- (e) I agree almost entirely with Mia.

Appendix 3B: EBAPS Scoring Scheme

EBAPS scoring scheme

Each item is scored on a scale of 0 (least sophisticated) to 4 (most sophisticated).

The [scoring scheme](#) is non-linear to take into account question-by-question variations in whether, for instance, neutrality is more or less sophisticated. A subscale score is simply the average of the student's scores on every item in that subscale. (When an item within a given subscale is left blank, the average is calculated without that item included.)

Sometimes we multiply through by 25 in order to report subscale scores on a scale of 0 to 100.

To automate the scoring using Microsoft Excel, download the Excel scoring [template](#).

A: Strongly disagree B: Somewhat disagree C: Neutral D: Somewhat agree E: Strongly agree

1. A = 4, B = 3, C = 1, D = 0.5, E = 0
2. A = 0, B = 1.5, C = 2.5, D = 3.5, E = 4
3. A = 0, B = 1, C = 2, D = 3.5, E = 4
4. A = 4, B = 3, C = 2, D = 1, E = 0
5. A = 0, B = 1, C = 2, D = 3, E = 4
6. A = 4, B = 4, C = 2, D = 1, E = 0
7. A = 4, B = 3, C = 2, D = 1, E = 0
8. A = 4, B = 3, C = 1.5, D = 0.5, E = 0
9. A = 0, B = 1, C = 2, D = 3, E = 4
10. A = 4, B = 3, C = 2, D = 1, E = 0
11. A = 0, B = 1, C = 2, D = 3, E = 4
12. A = 0, B = 0.5, C = 1, D = 3, E = 4
13. A = 4, B = 3, C = 1, D = 0.5, E = 0
14. A = 4, B = 3, C = 2, D = 1, E = 0
15. A = 4, B = 3, C = 2, D = 1, E = 0
16. A = 0, B = 1, C = 2, D = 3, E = 4
17. A = 4, B = 3, C = 1.5, D = 0.5, E = 0
18. A = 4, B = 3.5, C = 1.5, D = 0.5, E = 0
19. A = 4, B = 0, C = 3, D = 2, E = 1
20. A = 4, B = 0, C = 3, D = 2, E = 1
21. A = 4, B = 3, C = 2, D = 1, E = 0
22. A = 4, B = 3, C = 2, D = 1, E = 0
23. A = 0, B = 4, C = 1, D = 2, E = 3
24. A = 4, B = 4, C = 2, D = 1, E = 0
25. A = 0, B = 1, C = 2, D = 4, E = 4
26. A = 4, B = 4, C = 2, D = 1, E = 0
27. A = 4, B = 4, C = 2, D = 1, E = 0
28. A = 0, B = 1, C = 2, D = 3, E = 4
29. A = 0, B = 2, C = 4, D = 2, E = 0
30. A = 0, B = 1, C = 2, D = 3, E = 4

Appendix 4A: A3Home work Quiz

Question1

A capacitor with a charge of $[q]$ C has its terminals shorted by a metal wire so that the charge flows off within $[t]$ s. What is the average current flowing during that time in Ampere (A)?

Question2

A high voltage transmission line has an aluminum cable of diameter $[d]$ cm and is $[x]$ km long. The resistivity of Aluminum is $2.8 \times 10^{-8} \Omega\text{m}$. What is the resistance of this cable in Ω ?

Question3

A wire carries a steady current of $[x]$ A. What total charge (in C) passes through a cross-sectional area of the wire in a $[t]$ s time interval?

Question4

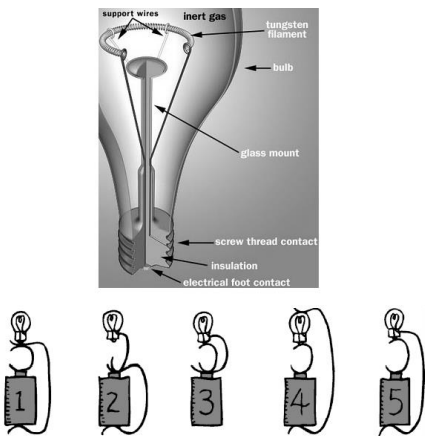
A current of $[x]$ A flows through the heating element of heater converting $[y]$ J of electrical energy into thermal energy every second. What is the voltage (in V) across the ends of the heating element?

Question5

Three resistors connected **in series** carry currents labeled I_1 , I_2 , I_3 , respectively. Which of the following expresses the value of the total current I_T in the system made up of the three resistors?

Question6

The figure below shows different ways to connect a light bulb to a battery with wires. Which of the bulbs will light up?



Question7

Two capacitors have an equivalent capacitance of 8 pF if connected in parallel, and 2 pF, if connected in series. Find C_1 and C_2 .

Question8

A parallel combination of resistors has an equivalent (or effective) resistance of 2Ω . Which one of the following is true?

Appendix 4B: Test 1

Question 11 pts

You have just pulled your clothes from the dryer and find that a sock is clinging to your jeans with static electricity. You pull the two apart. As the jeans and the sock move apart, the force between them becomes weaker because the

- ☐ the garments become polarized.
- ☐ electric charge on each garment diminishes as they move apart.
- ☐ electric charge on each garment increases as they move apart.
- ☐ electric current moving through each garment diminishes as they move apart.
- ☐ the force between the electric charges becomes weaker with increasing distance.

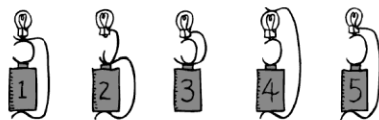
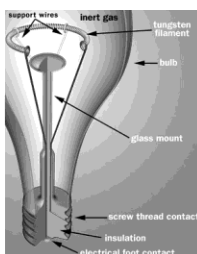
Question 21 pts

Which of the following statements is true?

- ☐ A positive charge experiences an attractive electrostatic force toward a nearby neutral conductor.
- ☐ A positive charge experiences no electrostatic force near a neutral conductor.
- ☐ A positive charge experiences a repulsive force, away from a nearby conductor.
- ☐ Whatever the force on a positive charge near a neutral conductor, the force on a negative charge is then oppositely directed.

Question 31 pts

The figure below shows different ways to connect a light bulb to a battery with wires. Which of the bulbs will light up?



- ☐ 4
- ☐ 1
- ☐ 2

- ☐ 3
- ☐ 5

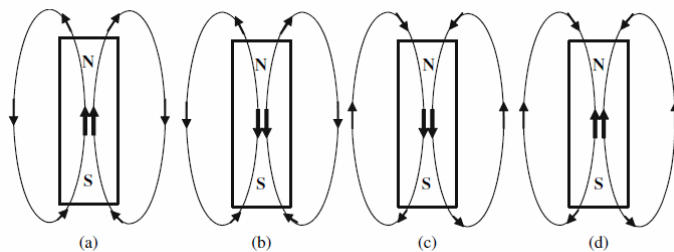
Question 41 pts

You double the voltage across a certain conductor and you observe the current increases three times. Does the conductor obey Ohm's law?

- ☐ Ohm's law states that the resistance increases with temperature. So not enough information is given to decide if the conductor obeys Ohm's law.
- ☐ The conductor obeys Ohm's law since the current increases when V increases.
- ☐ The conductor does not obey Ohm's law.

Question 51 pts

Which one of the following best illustrates magnetic field lines associated with the bar magnets shown?



- ☐ (c)
- ☐ (a)
- ☐ (d)
- ☐ (b)
- ☐ None of these, since the field lines encircle the bar magnet according to the right-hand rule.

Question 61 pts

A current in a long, straight wire produces a magnetic field. The magnetic field lines

- ☐ come in from infinity to the wire.
- ☐ form circles that go around the wire.
- ☐ go out from the wire to infinity.
- ☐ form circles that pass through the wire.
- ☐ are parallel to the wire.

Question 71 pts

Find the magnitude of the electric field produced by a point charge $q = 2.37$ nanoCoulomb a distance $r = 2.7$ meters from the charge in units of N/C .

Enter a number with two digits behind the decimal point.

Question 81 pts

Two 1 microCoulomb point charges are 2 m apart. The potential at each charge due to the other charge is

- ☐ 2250 V.
- ☐ 9000 V,
- ☐ 4500 V,
- ☐ 1125 V.
- ☐ 0.0045 V.

Question 91 pts

If 500 J of work are required to carry a 40 C charge from one point to another, the potential difference between these two points is

- ☐ 20000 V
- ☐ undefined. It depends on the path the charged was carried on.
- ☐ 12.5 V
- ☐ 0.08 V

Question 101 pts

A mass spectrometer is to be used to separate protons from deuterium nuclei. A deuterium nucleus has the same charge and twice the mass as a proton, since it contains an extra neutron. Both the deuterium and the proton nuclei are accelerated by the same voltage. Which of the following statements is true?

- ☐ Both the proton and the deuterium will move along the same circular path.
- ☐ The deuterium will have a larger radius of curvature, since it is more massive.
- ☐ The deuterium will have a smaller radius of curvature, since it is more massive.
- ☐ The deuterium will not be deflected in the mass spectrometer, since it contains a neutron.
- ☐ Mass spectrometers cannot be used to analyze nuclei, only molecular ions.

Question 111 pts

Can an electron **move** through a region of magnetic field without being deflected?

- ☐ No.
- ☐ Yes, but only if it travels in a circle.
- ☐ Yes, if it travels in a direction parallel or anti-parallel to the magnetic field.
- ☐ Yes, but only if it travels in a direction parallel to the magnetic field.
- ☐ Yes, if it travels perpendicular to the magnetic field.

Question 121 pts

Which of the following statements about a conductor in electrostatic equilibrium is **false**?

- (A) The electric field inside the conductor is zero.
- (B) Just outside the conductor, the electrostatic field is perpendicular to its surface.
- (C) The net charge inside the conductor is zero, all excess charge resides on the surface.
- (D) A charge located within a hole in a conductor at equilibrium feels no force from charges outside the conductor.

- ☐ Statement (C)
- ☐ Statement (A)
- ☐ Statement (B)
- ☐ All the statements are true.
- ☐ Statement (D)

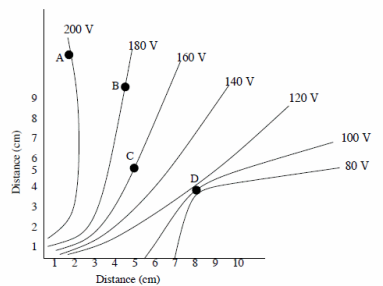
Question 131 pts

A cubical closed surface has N inward electric flux lines passing through one face and N outward electric flux lines passing through the opposite face. No flux lines are passing through any other face. If the magnitude of the electric field at the center of the cube is E , the total charge inside the cube is

- ☐ $-\epsilon_0 N$.
- ☐ zero.
- ☐ $\epsilon_0 N$.
- ☐ $-2\epsilon_0 N$.
- ☐ $2\epsilon_0 N$.

Question 141 pts

Equipotential lines are shown. For which region is the magnitude of the electric field the greatest?



- ☐ It is the same everywhere.
- ☐ B
- ☐ C

- ☐ A
- ☐ D

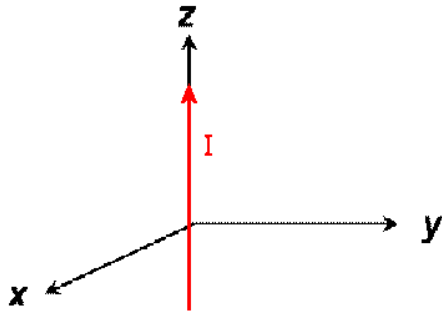
Question 151 pts

In the early universe and in stars, deuterium nuclei are produced from the combination of **one proton** and **one neutron**, with the release of a gamma ray. What is the charge on a deuterium nucleus in units of $q_e = 1.6 \times 10^{-19} \text{ C}$?

- ☐ 0
- ☐ -2
- ☐ 2
- ☐ 1
- ☐ -1

Question 161 pts

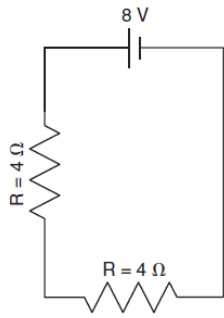
A long straight vertical wire carries an 8 A current upwards along the z axis. Find the magnetic field produced by this wire (magnitude and direction) on the x-axis at $x = 4 \text{ m}$.



- ☐ 8 microTesla in the negative y-direction
- ☐ 1.6 microTesla in the negative x-direction
- ☐ 1.6 microTesla in the positive x-direction
- ☐ 0.4 microTesla in the positive y-direction
- ☐ 8 microTesla in the positive y-direction
- ☐ 0.4 Tesla in the negative y-direction

Question 171 pts

What is the current through the battery in the circuit shown below?



- ☐ 64 A
- ☐ 16 A
- ☐ 4 A
- ☐ 1 A
- ☐ 8 A

Appendix 5: T-tests and Network Graphs Experiment 1

Table A5. 1 T-test for Significance of Differences in Epistemic Networks by EBAPS Scores - Low, Medium and High (PreQ and Post Q)

	PreQ Dim 1	Dim 2	Post Q Dim1	Dim 2
Mean Low	-0.382	-0.15	0.405	0.001
Mean Medium	-0.418	-0.03	0.418	0.012
Mean High	-0.51	0.144	0.373	-0.069
Low-Medium				
T statistic	0.391	-0.559	-0.095	-0.127
P value	.71	.598	.926	.9
Cohen's d	0.223	-0.323	-0.04	-0.053
Medium-High				
T statistic	1.575	-1.325	0.379	1.008
P value	.127	.196	.708	.322
Cohen's d	0.575	-0.481	0.138	0.368
Low-High				
T statistic	1.35	-1.336	0.235	0.8
P value	.222	.229	.817	.433
Cohen's d	0.782	-0.801	0.103	0.33

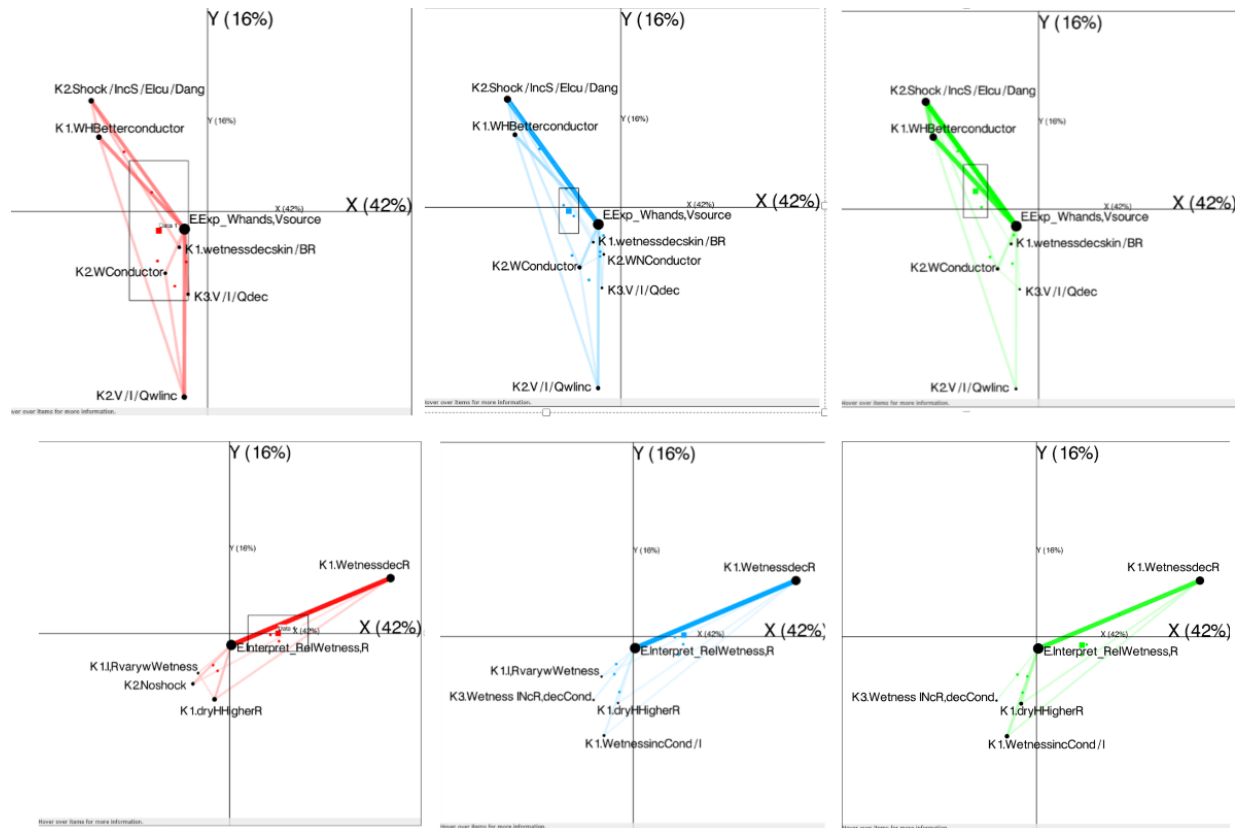


Figure A5. 1. Mean networks by EBAPS overall scores - Low (red), Medium (blue), and High (green), for PreQ (Upper 3 graphs), and Post Q (Lower 3 graphs).

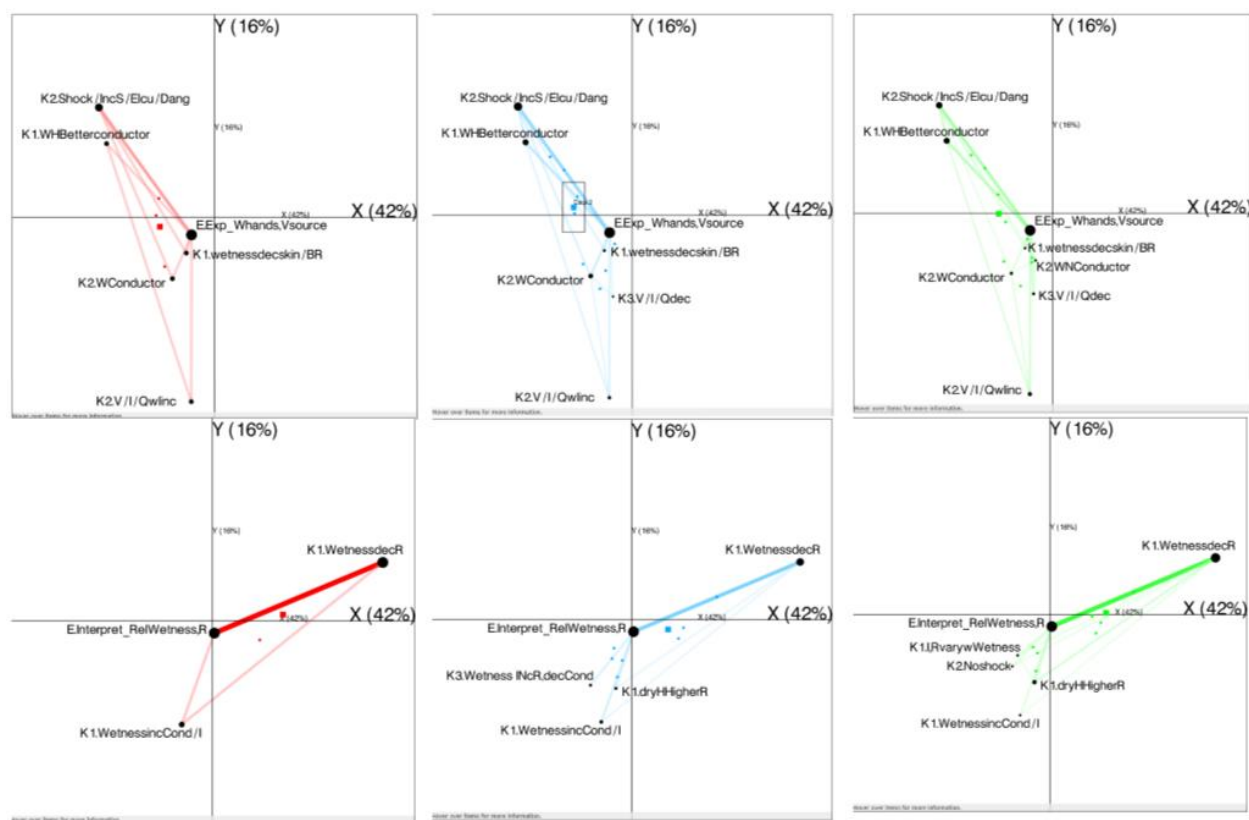


Figure A5. 2. Mean networks by A3HW Quiz (Attempt1) Scores- Low (red), Medium (blue), and High (green), for PreQ (Upper 3 graphs), and Post Q (Lower 3 graphs).

Table A5. 2 T-test for Significance of Differences in Epistemic Networks by A3HWQuiz Scores - Low, Medium and High (PreQ and Post Q)

	PreQ		Post Q	
	Dim 1	Dim 2	Dim1	Dim 2
Mean Low	-0.43	-0.074	0.557	0.049
Mean Medium	-0.467	0.063	0.29	-0.078
Mean High	-0.433	-0.004	0.455	0.014
Low-Medium				
T statistic	0.793	-0.736	2.158	1.231
P value	0.439	0.508	0.046	0.248
Cohen's d	0.247	-0.399	0.785	0.554
Medium-High				
T statistic	-0.561	0.505	-1.431	-1.21
P value	0.579	0.617	0.164	0.237
Cohen's d	-0.194	0.175	-0.512	-0.429
Low-High				
T statistic	-2.168	-0.448	1.056	0.379
P value	0.04	0.684	0.316	0.717
Cohen's d	-0.491	-0.223	0.405	0.189

Table A5. 3 T-test for Significance of Differences in Epistemic Networks by Lab4Quiz Scores - Low, Medium and High (PreQ and Post Q)

	PreQ		Post Q	
	Dim 1	Dim 2	Dim1	Dim 2
Mean Low	-0.42	-0.07	0.469	0.063
Mean Medium	-0.441	0.055	0.343	-0.081
Mean High	-0.483	0.075	0.421	-0.013
Medium-Low				
T statistic	-0.605	0.129	-1.002	-1.793
P value	0.551	0.899	0.326	0.084
Cohen's d	-0.253	0.054	-0.384	-0.666
High-Medium				
T statistic	-0.561	0.505	0.662	0.844
P value	0.579	0.617	0.514	0.406
Cohen's d	-0.194	0.175	0.252	0.317
High-Low				
T statistic	-0.907	0.923	-0.379	-0.966
P value	0.374	0.366	0.708	0.344
Cohen's d	-0.379	0.385	-0.157	-0.403

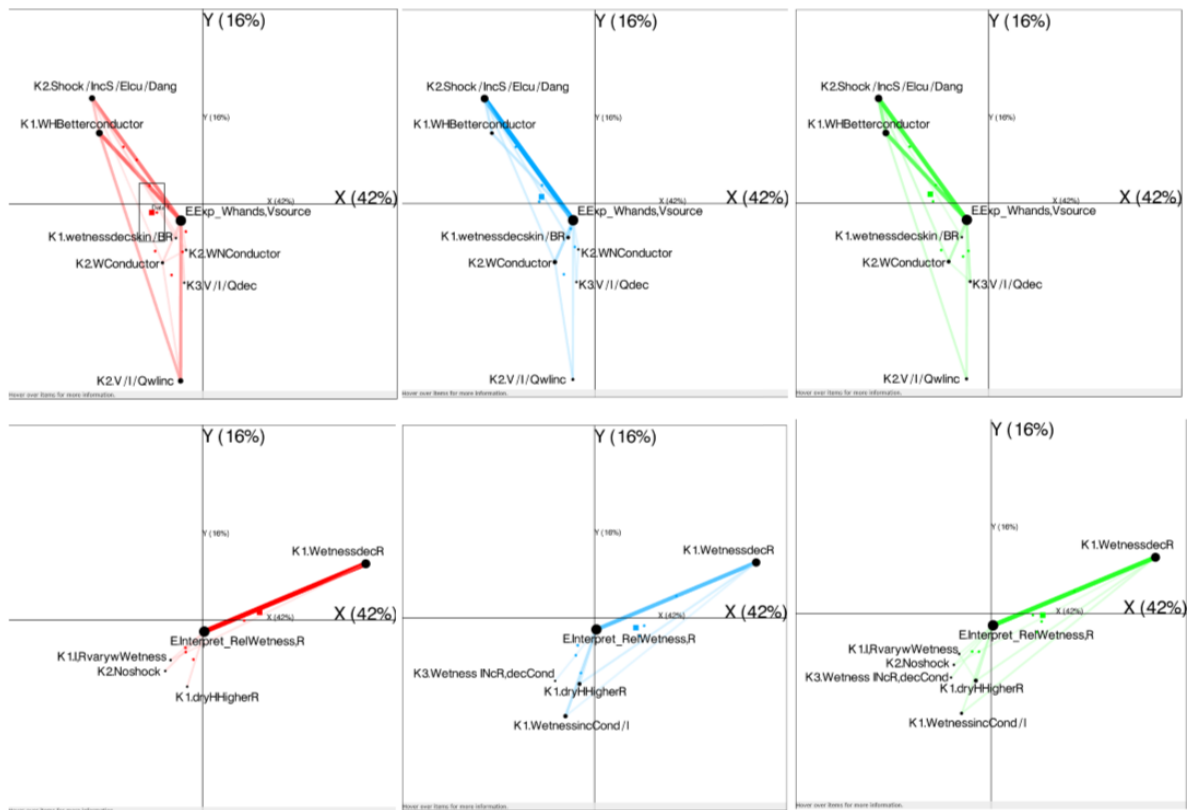


Figure A5. 3. Mean networks by Lab4 Quiz Scores - Low (red), Medium (blue), and High (green), for PreQ (Upper 3 graphs), and Post Q (Lower 3 graphs).

Table A5. 4 T-test for Significance of Differences in Epistemic Networks by Lab4 Scores - Low, Medium and High (PreQ and Post Q)

	PreQ		Post Q	
	Dim 1	Dim 2	Dim1	Dim 2
Mean Low	-0.416	-0.057	0.388	-0.017
Mean Medium	-0.457	0.019	0.42	-0.016
Mean High	-0.416	-0.057	0.402	-0.018
Medium-Low				
T statistic	-0.658	0.522	0.236	0.011
P value	0.517	0.607	0.816	0.991
Cohen's d	-0.265	0.21	0.097	0.005
High-Medium				
T statistic	0.658	-0.522	-0.153	-0.025
P value	0.517	0.607	0.88	0.981
Cohen's d	0.265	-0.21	-0.058	-0.009
High-Low				
T statistic	0.781	-1.079	0.109	-0.008
P value	0.446	0.294	0.915	0.993
Cohen's d	0.343	-0.465	0.045	-0.004

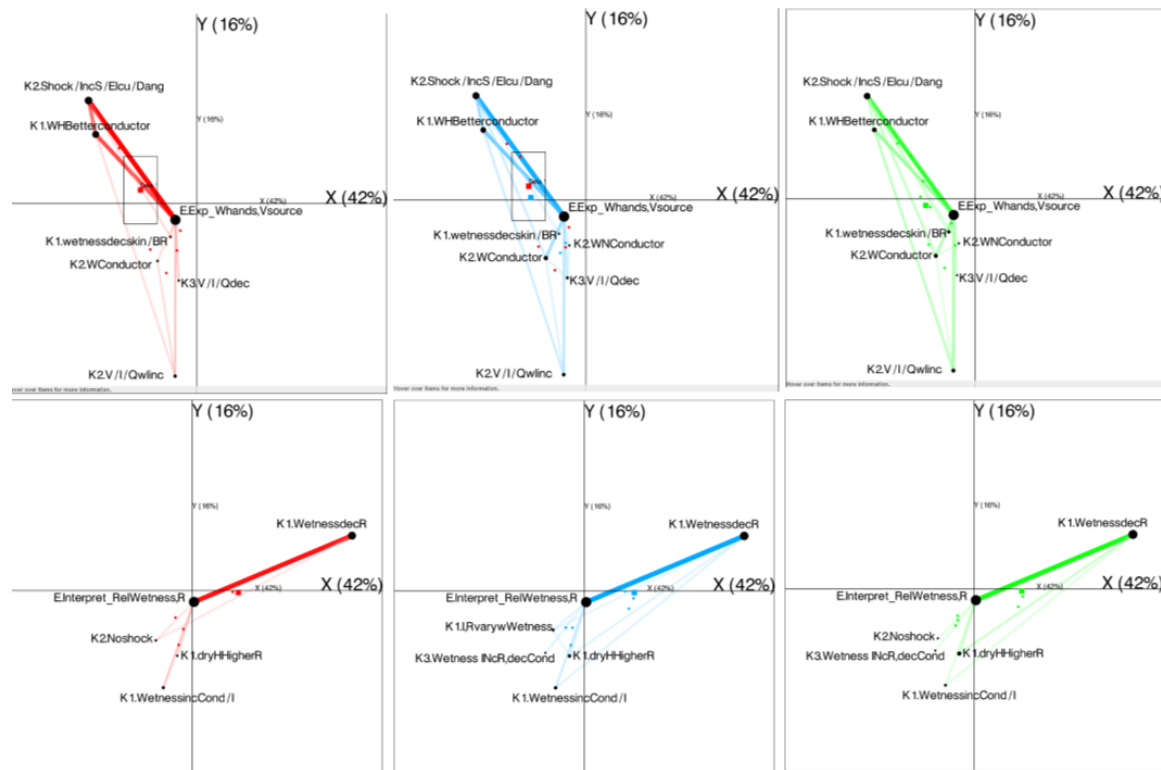


Figure A5. 4. Mean networks by Lab4 Scores - Low (red), Medium (blue), and High (green), for PreQ (Upper 3 graphs), and Post Q (Lower 3 graphs).

Table A5. 5 T-test for Significance of Differences in Epistemic Networks by Test1 Scores - Low, Medium and High (PreQ and Post Q)

	PreQ		Post Q	
	Dim 1	Dim 2	Dim1	Dim 2
Mean Low	-0.428	-0.011	0.388	-0.004
Mean Medium	-0.498	0.098	0.473	0.009
Mean High	-0.387	-0.073	0.296	-0.077
Medium-Low				
T statistic	-1.032	0.709	0.616	0.14
P value	0.316	0.486	0.546	0.89
Cohen's d	-0.437	0.287	0.264	0.057
High-Medium				
T statistic	0.658	-0.522	-1.6	-1.214
P value	0.517	0.607	0.125	0.236
Cohen's d	0.265	-0.21	-0.628	-0.434
High-Low				
T statistic	1.714	-1.134	-0.608	-0.773
P value	0.107	0.271	0.55	0.45
Cohen's d	0.778	-0.479	-0.272	-0.346

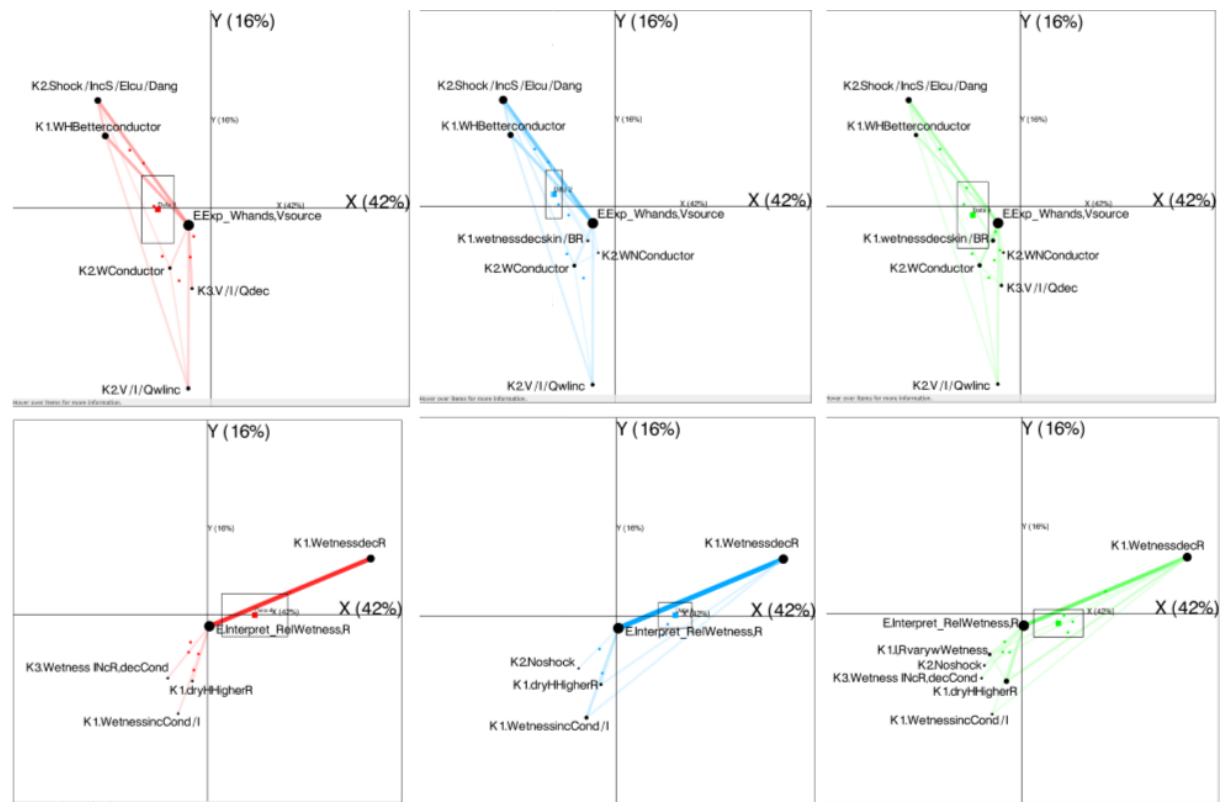


Figure A5. 5 . Mean networks by Test1 Scores - Low (red), Medium (blue), and High (green), for PreQ (Upper 3 graphs), and Post Q (Lower 3 graphs).

Appendix 6: T-tests and Network Graphs Experiment 2

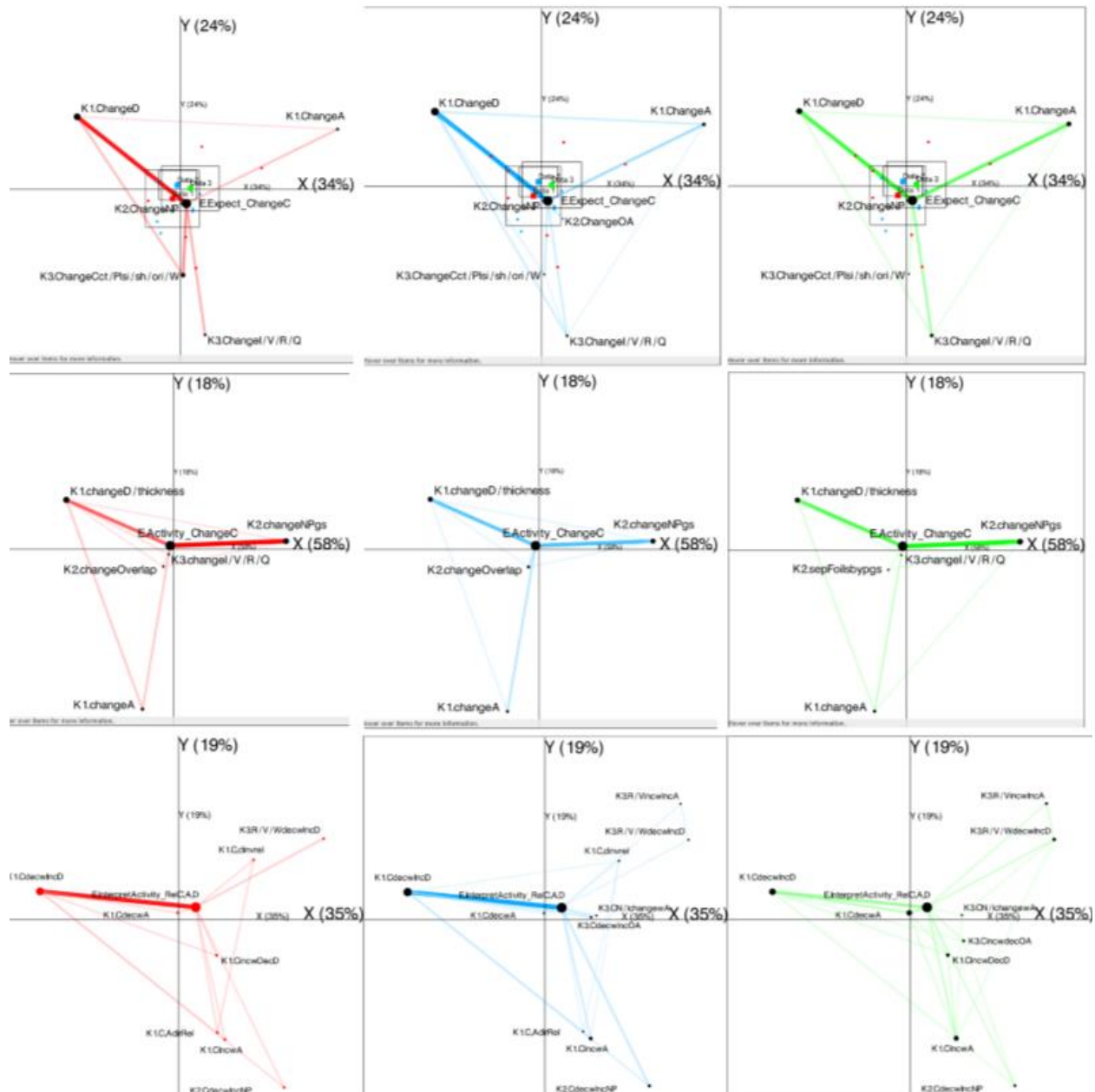


Figure A6. 1. Mean Network graphs of students grouped by EBAPS scores, Low (red), Medium (blue), High (green), for PreQ (Upper 3 graphs), Post Q1(center 3 graphs), and PostQ2(Lower 3 graphs).

Table A6. 1 T-test for Significance of Differences in Epistemic Networks of students grouped by EBAPS scores, Low, Medium, and High. [PreQ, Post Q1, and PostQ2]

	PreQ Dim 1	Dim 2	PostQ1 Dim1	Dim 2	PostQ2 Dim1	Dim2
Mean Low	-0.086	-0.098	0.039	0.001	-0.091	0.023
Mean Medium	-0.033	0.046	-0.036	-0.02	-0.113	-0.04
Mean High	0.106	0.013	0.021	0.03	0.116	-0.007
Medium-Low						
T statistic	0.331	0.929	-0.292	-0.148	-0.128	-0.535
P value	0.744	0.367	0.774	0.883	0.9	0.599
Cohen's d	0.13	0.4	-0.131	-0.059	-0.058	-0.213
High-Medium						
0.258	0.389	-0.247	0.258	0.389	1.53	0.276
0.798	0.701	0.807	0.798	0.701	0.137	0.784
0.096	0.143	-0.091	0.096	0.143	0.569	0.102
High-Low						
T statistic	1.005	0.632	-0.067	0.232	1.133	-0.256
P value	0.325	0.534	0.948	0.819	0.275	0.801
Cohen's d	0.399	0.263	-0.029	0.102	0.514	-0.105

Table A6. 2 T-test for Significance of Differences in Epistemic Networks of students grouped by A3HWQuiz scores, Low, Medium, and High. [PreQ, Post Q1, and PostQ2]

	PreQ Dim 1	Dim 2	Post Q1 Dim1	Dim 2	PostQ2 Dim1	Dim2
Mean Low	0.259	-0.113	-0.232	-0.006	0.306	-0.095
Mean Medium	0.013	-0.001	0.033	0.079	0.063	-0.025
Mean High	-0.066	0.025	0.032	-0.055	-0.13	0.044
Medium-Low						
T statistic	-0.859	0.502	0.864	0.527	-0.99	0.294
P value	0.424	0.636	0.412	0.617	0.358	0.78
Cohen's d	-0.507	0.329	0.425	0.317	-0.564	0.191
High-Medium						
T statistic	-0.543	0.218	-0.003	-1.22	-1.327	0.638
P value	0.591	0.829	0.998	0.231	0.195	0.53
Cohen's d	-0.183	0.069	-0.001	-0.4	-0.481	0.235
High-Low						
T statistic	-1.186	0.607	0.948	-0.286	-0.13	0.044
P value	0.29	0.568	0.38	0.783	0.306	-0.095
Cohen's d	-0.76	0.331	0.496	-0.133	-1.858	0.609

Table A6. 3. *T-test for Significance of Differences in Epistemic Networks of students grouped by Lab4Quiz scores, Low, Medium, and High. [PreQ, Post Q1, and PostQ2]*

	PreQ Dim 1	Dim 2	Post Q1 Dim1	Dim 2	PostQ2 Dim1	Dim2
Mean Low	-0.204	0.156	0.078	-0.004	0.103	0.056
Mean Medium	0.017	-0.044	0.057	0.008	-0.096	-0.1
Mean High	0.173	-0.099	-0.161	-0.006	-0.02	0.055
Medium-Low						
T statistic	1.442	-1.508	-0.094	0.104	-1.228	-1.229
P value	0.16	0.144	0.926	0.918	0.23	0.23
Cohen's d	0.544	-0.556	-0.036	0.041	-0.469	-0.47
High-Medium						
T statistic	0.933	-0.37	-0.924	-0.105	0.434	1.383
P value	0.359	0.714	0.365	0.917	0.67	0.181
Cohen's d	0.348	-0.137	-0.37	-0.044	0.192	0.575
High-Low						
T statistic	2.254	-1.993	-0.982	-0.013	-0.66	-0.008
P value	0.033	0.057	0.337	0.99	0.517	0.994
Cohen's d	0.862	-0.759	-0.412	-0.005	-0.275	-0.003

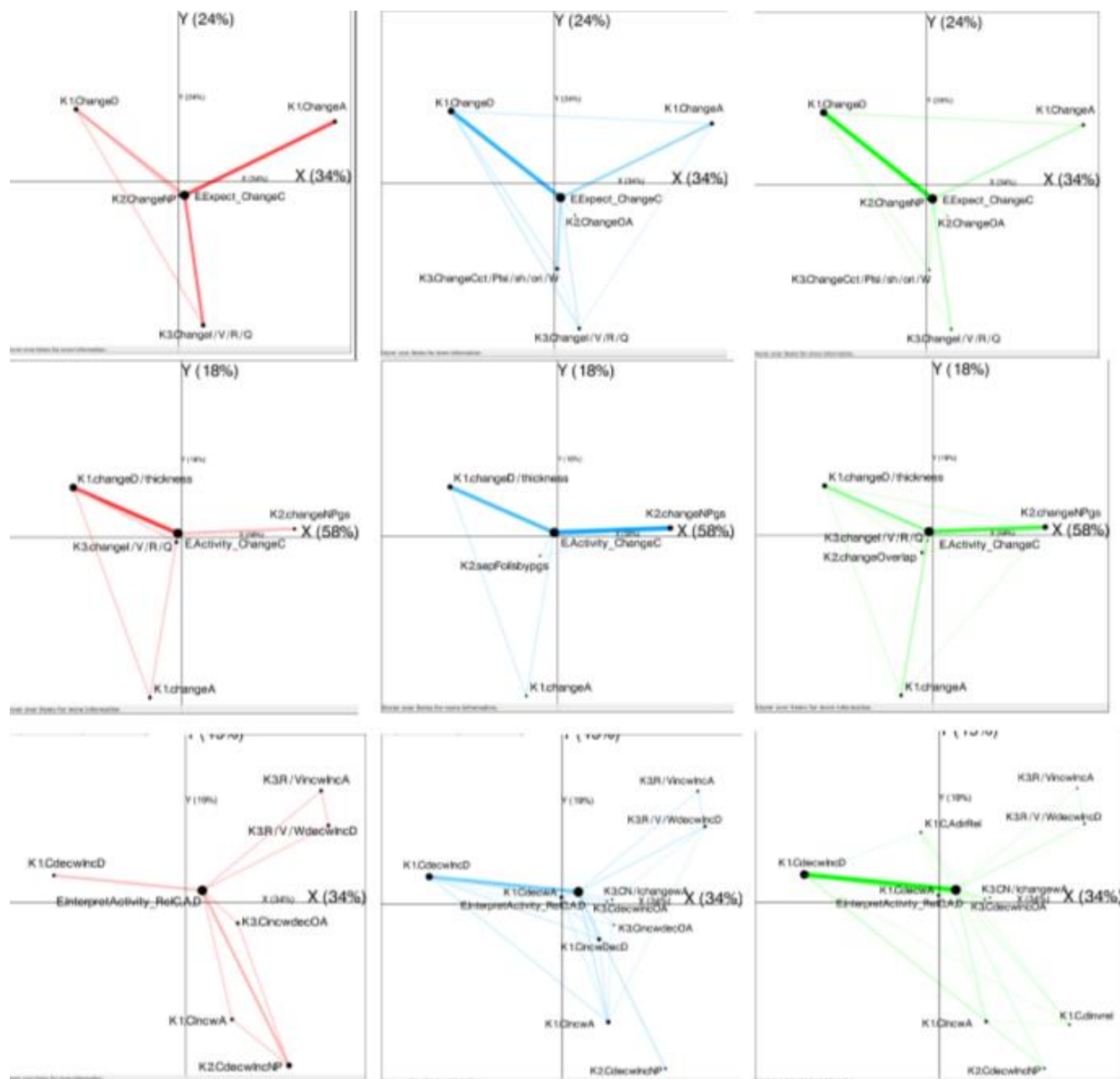


Figure A6. 2. Mean Network graphs of students grouped by A3HWQuiz scores, Low (red), Medium (blue), High (green), for PreQ (Upper 3 graphs), Post Q1(center 3 graphs), and PostQ2(Lower 3 graphs).

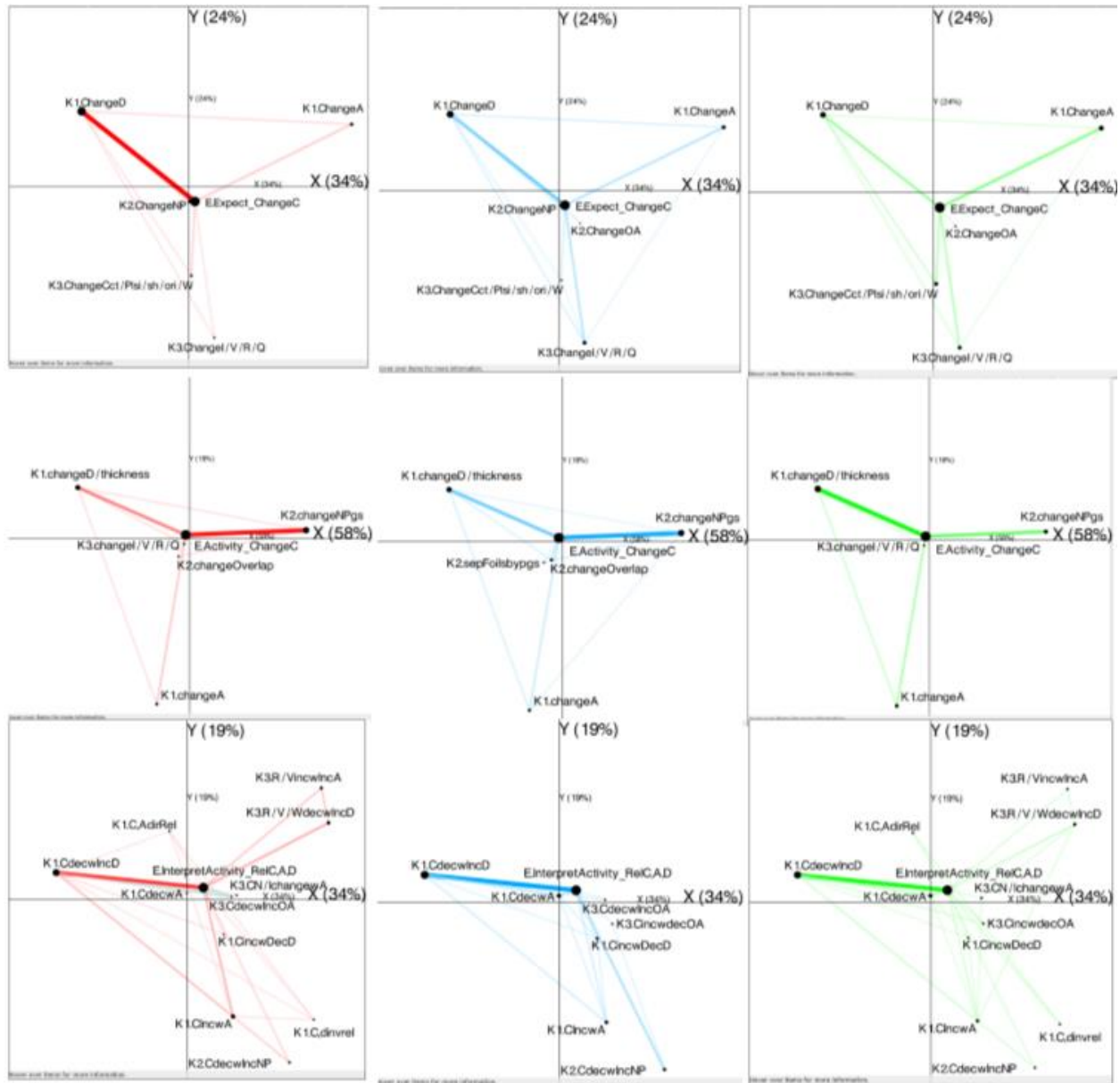


Figure A6.3. Mean Network graphs of students grouped by Lab4Quiz scores, Low (red), Medium (blue), High (green), for PreQ (Upper 3 graphs), Post Q1(center 3 graphs), and PostQ2(Lower 3 graphs).

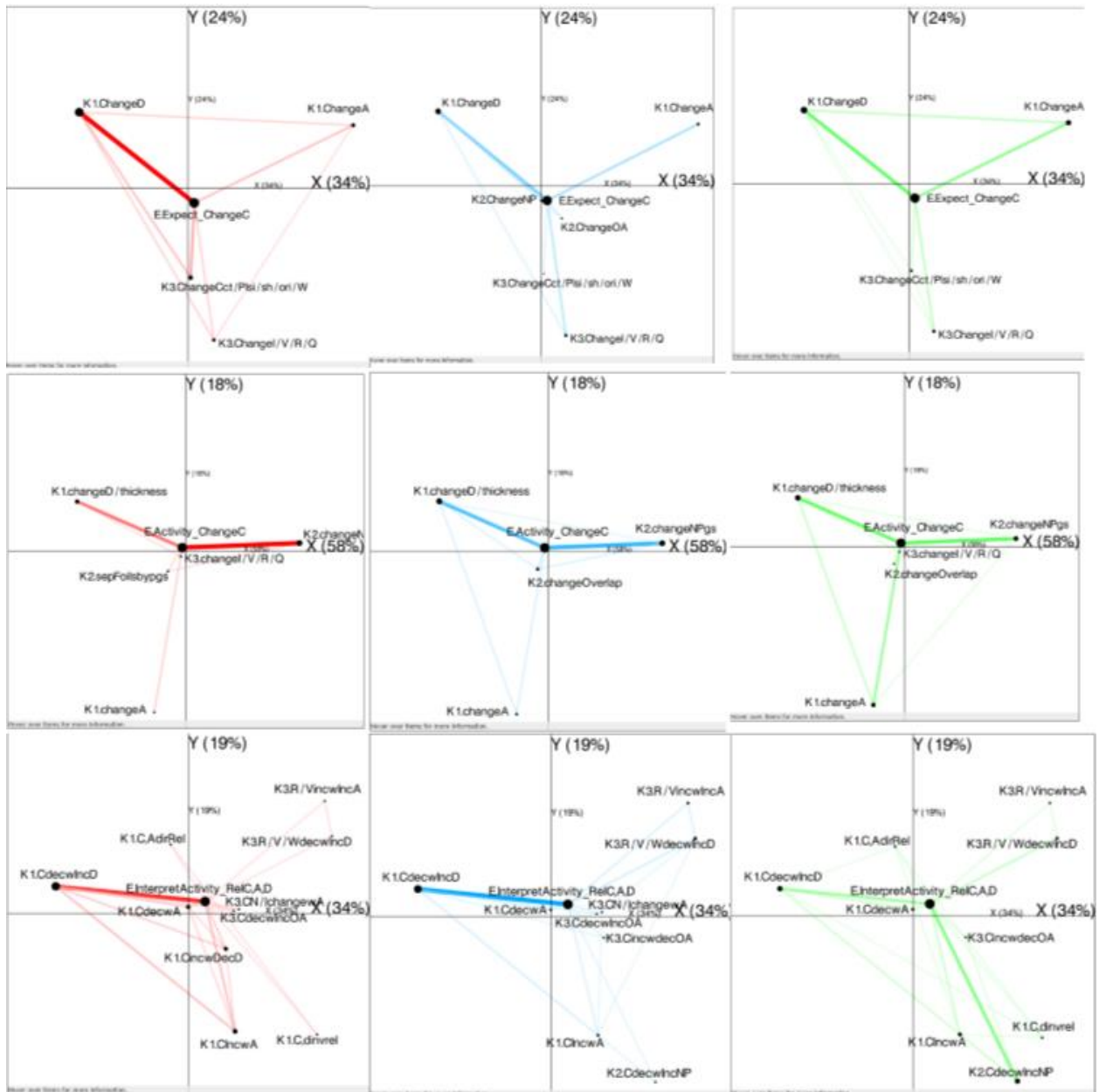


Figure A6. 4Mean Network graphs of students grouped by Lab4 scores, Low (red), Medium (blue), High (green), for PreQ (Upper 3 graphs), Post Q1(center 3 graphs), and PostQ2(Lower 3 graphs).

Table A6. 4. *T-test for Significance of Differences in Epistemic Networks of students grouped by Lab4 scores, Low, Medium, and High. [PreQ, Post Q1, and PostQ2]*

	PreQ Dim 1	Dim 2	Post Q1 Dim1	Dim 2	PostQ2 Dim1	Dim2
Mean Low	-0.154	0.035	0.09	0.056	-0.094	0.007
Mean Medium	0.046	-0.05	-0.004	0.062	-0.081	0.116
Mean High	0.173	-0.099	-0.066	-0.115	0.222	-0.167
Medium-Low						
T statistic	1.212	-0.706	-0.382	0.049	0.088	1.239
P value	0.235	0.486	0.706	0.962	0.931	0.226
Cohen's d	0.442	-0.256	-0.157	0.021	0.034	0.487
High-Medium						
T statistic	0.3	0.471	-0.283	-1.422	1.679	-1.899
P value	0.766	0.642	0.779	0.169	0.107	0.078
Cohen's d	0.11	0.18	-0.107	-0.554	0.687	-0.86
High-Low						
T statistic	1.574	-0.064	-0.636	-1.14	1.924	-1.153
P value	0.128	0.95	0.532	0.266	0.07	0.268
Cohen's d	0.618	-0.025	-0.271	-0.471	0.841	-0.519

Table A6. 5. *T-test for Significance of Differences in Epistemic Networks of students grouped by Test1 scores, Low, Medium, and High. [PreQ, Post Q1, and PostQ2]*

	PreQ Dim 1	Dim 2	Post Q1 Dim1	Dim 2	PostQ2 Dim1	Dim2
Mean Low	-0.084	-0.002	0.037	0.108	-0.037	-0.047
Mean Medium	0.044	-0.023	0.007	-0.008	0.024	-0.044
Mean High	0.039	0.043	-0.049	-0.095	0.001	0.151
Medium-Low						
T statistic	0.762	-0.164	-0.12	-1.059	0.392	0.024
P value	0.452	0.871	0.906	0.299	0.699	0.981
Cohen's d	0.268	-0.056	-0.048	-0.378	0.152	0.01
High-Medium						
T statistic	-0.03	0.401	-0.272	-0.615	-0.116	1.622
P value	0.976	0.692	0.788	0.546	0.909	0.122
Cohen's d	-0.011	0.155	-0.1	-0.248	-0.054	0.663
High-Low						
T statistic	0.733	0.303	-0.338	-1.471	0.174	1.424
P value	0.471	0.766	0.74	0.16	0.865	0.171
Cohen's d	0.294	0.132	-0.151	-0.658	0.083	0.639

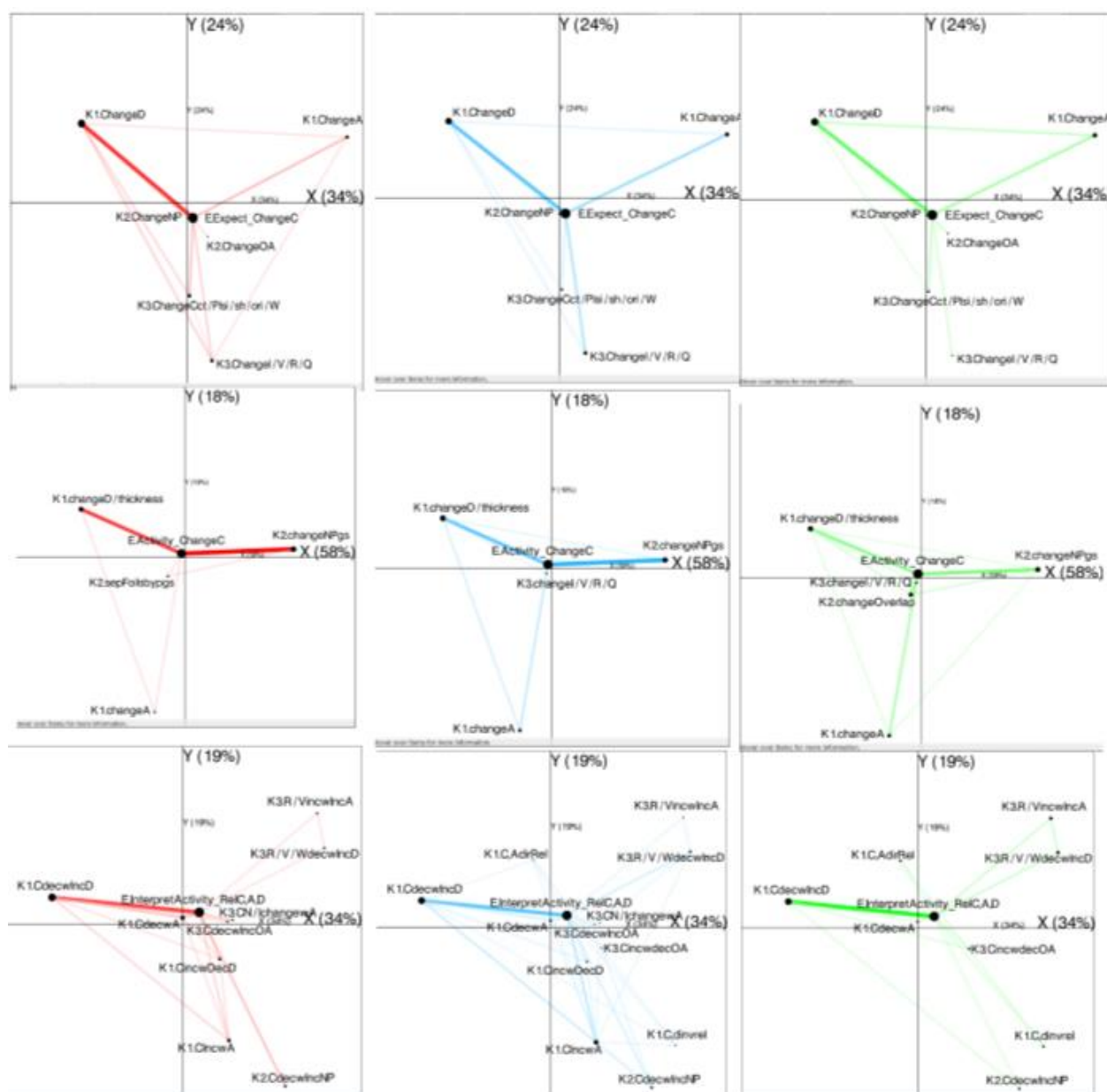


Figure A6. 5. Mean Network graphs of students grouped by Test1 scores, Low (red), Medium (blue), High (green), for PreQ (Upper 3 graphs), Post Q1(center 3 graphs), and PostQ2(Lower 3 graphs).

Appendix 7: T-tests and Network Graphs Experiment 3

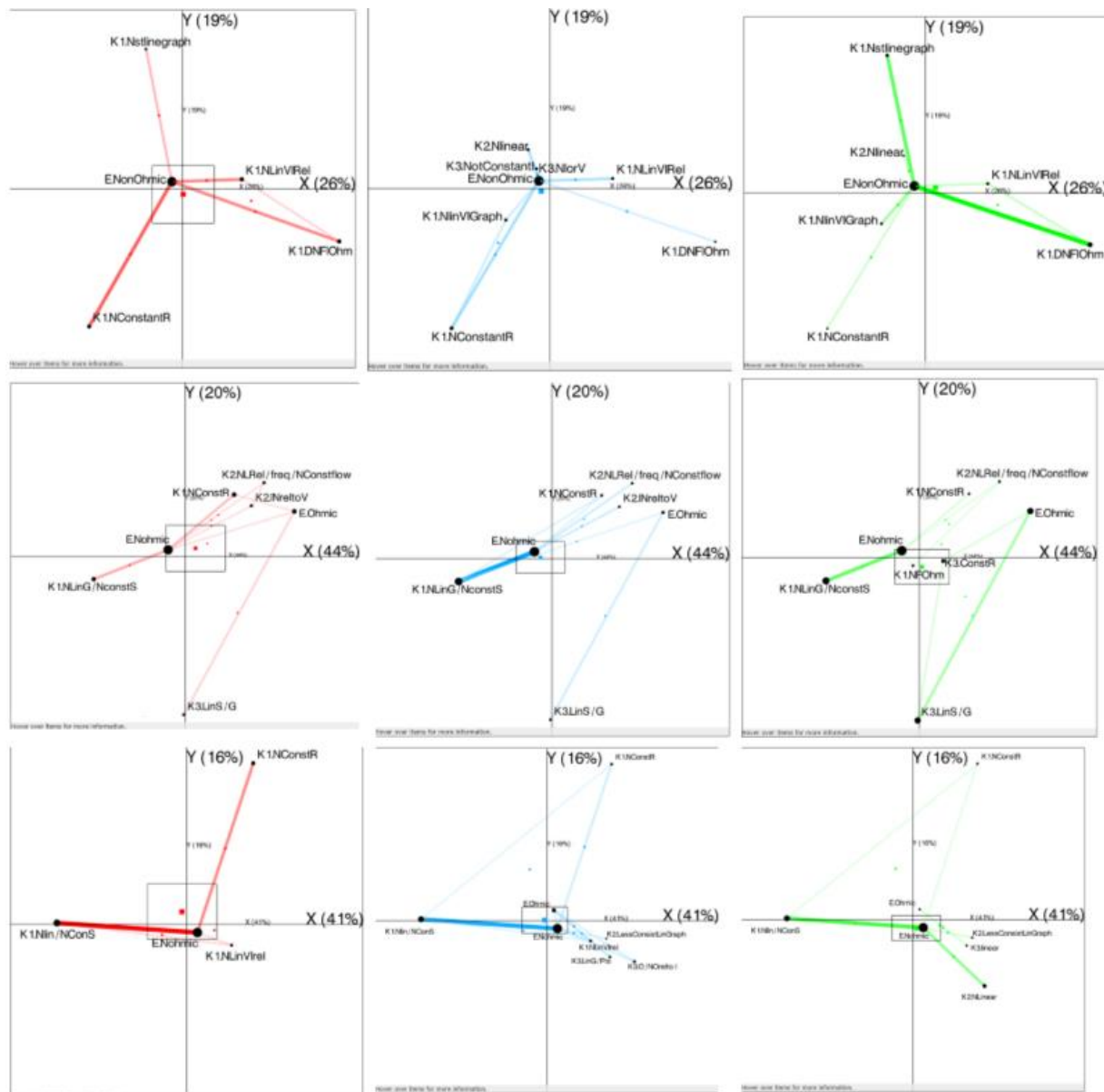


Figure A7. 1. Mean Network models of students by Low (red), Medium (blue), and High (green) EBAPS scores, for PreQ2 (upper 3graphs), PostQ2(center 3 graphs), and PostQ2 (lower 3 graphs).

Table A7. 1 T-test for Significance of Differences in Epistemic Networks of students by Low, Medium, and High EBAPS scores [PreQ2, Post Q1, and PostQ2]

	PreQ2		Post Q1		PostQ2	
	Dim 1	Dim 2	Dim1	Dim 2	Dim1	Dim2
Mean Low	0.008	-0.049	0.122	0.099	-0.051	0.14
Mean Medium	-0.08	-0.026	-0.133	0.013	-0.025	0.007
Mean High	0.039	0.043	0.034	-0.104	0.04	-0.094
Medium-Low						
T statistic	-0.59	0.171	-1.264	-0.583	0.126	-0.884
P value	0.564	0.867	0.222	0.568	0.901	0.393
Cohen's d	-0.27	0.078	-0.525	-0.256	0.055	-0.423
High-Medium						
T statistic	1.282	0.688	0.844	-0.919	0.368	-1.05
P value	0.212	0.498	0.406	0.366	0.716	0.303
Cohen's d	0.482	0.258	0.321	-0.349	0.139	-0.394
High-Low						
T statistic	0.485	0.669	-0.417	-1.339	0.432	-1.583
P value	0.633	0.511	0.681	0.198	0.671	0.139
Cohen's d	0.203	0.285	-0.179	-0.601	0.194	-0.783

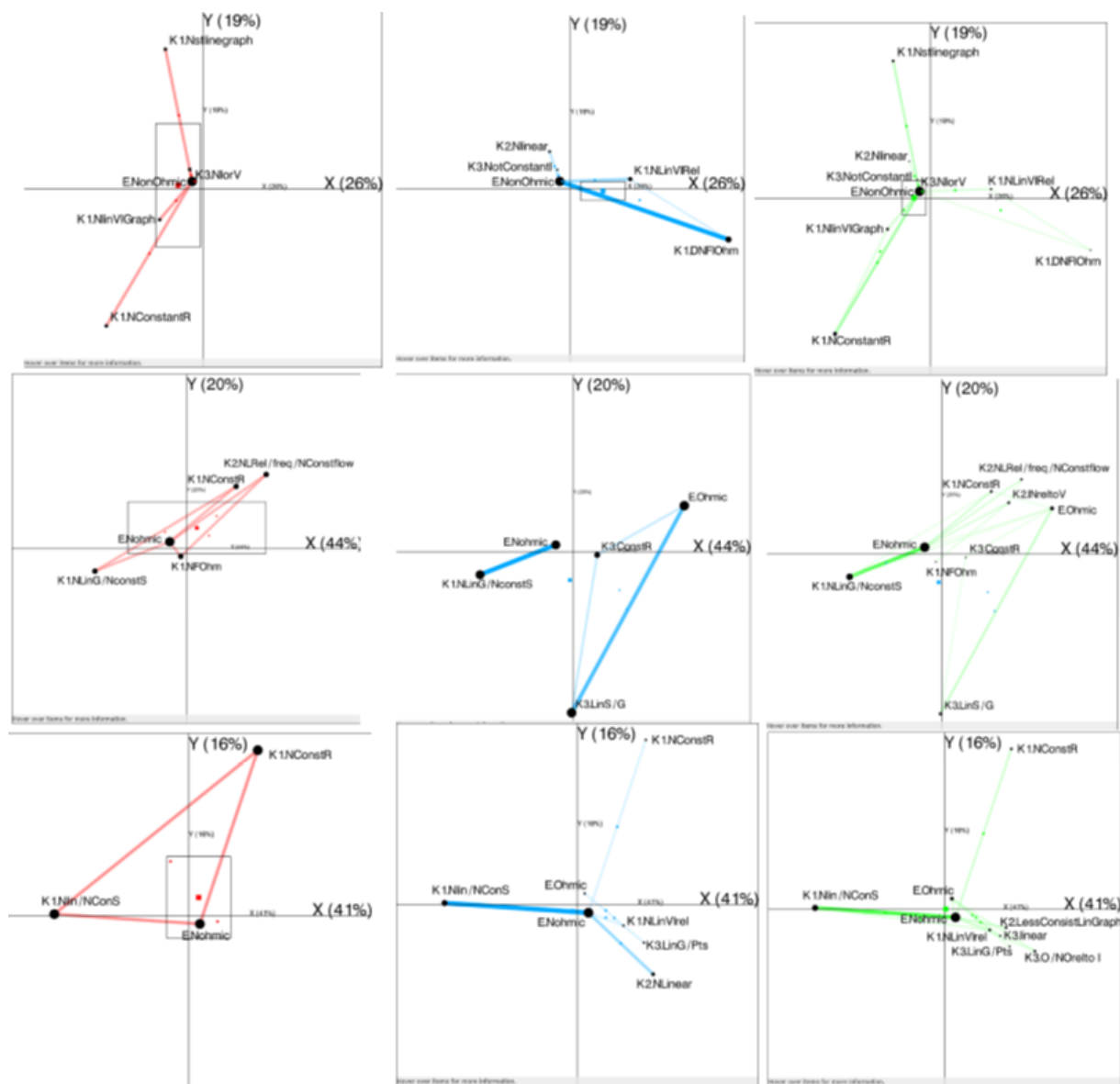


Figure A7. 2. Mean Network models of students by Low (red), Medium (blue), and High (green) A3HW Quiz scores, for PreQ2 (upper 3graphs), PostQ2(center 3 graphs), and PostQ2 (lower 3 graphs).

Table A7. 2 T-test for Significance of Differences in Epistemic Networks of students by Low, Medium, and High A3HW Quiz scores [PreQ2, Post Q1, and PostQ2]

	PreQ2		Post Q1		PostQ2	
	Dim 1	Dim 2	Dim1	Dim 2	Dim1	Dim2
Mean Low	-0.233	0.033	0.117	0.245	0.107	0.194
Mean Medium	0.314	-0.023	-0.038	-0.329	-0.056	-0.075
Mean High	-0.154	0.008	-0.014	0.013	0.012	0.004
Medium-Low						
T statistic	4.376	-0.265	-0.385	-3.686	-0.858	-1.52
P value	<.001	0.803	0.718	0.018	0.405	0.177
Cohen's d	1.642	-0.226	-0.312	-2.975	-0.35	-0.902
High-Medium						
T statistic	-4.106	0.351	0.066	2.266	0.384	0.783
P value	<.001	0.728	0.951	0.083	0.705	0.441
Cohen's d	-1.521	0.103	0.047	1.038	0.14	0.283
High-Low						
T statistic	-0.233	0.033	-0.628	-2.497	0.432	-1.583
P value	0.863	-0.111	0.575	0.049	0.671	0.139
Cohen's d	0.411	0.915	-0.267	-0.713		

Table A7. 3 T-test for Significance of Differences in Epistemic Networks of students by Low, Medium, and High Lab4 Quiz scores [PreQ2, Post Q1, and PostQ2]

	PreQ2		Post Q1		PostQ2	
	Dim 1	Dim 2	Dim1	Dim 2	Dim1	Dim2
Mean Low	-0.093	-0.136	0.065	0.216	0.107	0.194
Mean Medium	-0.008	0.007	0.156	-0.12	-0.056	-0.075
Mean High	-0.154	0.008	-0.279	-0.125	0.012	0.004
Medium-Low						
T statistic	0.574	1.149	0.523	-2.836	0.77	-1.267
P value	0.57	0.261	0.605	0.01	0.449	0.217
Cohen's d	0.221	0.441	0.205	-1.112	0.302	-0.497
High-Medium						
T statistic	0.806	1.056	-2.076	-0.033	-3.569	-0.491
P value	0.427	0.301	0.05	0.974	0.002	0.63
Cohen's d	0.309	0.402	-0.859	-0.013	-1.503	-0.189
High-Low						
T statistic	1.38	2.382	-1.742	-3.252	-2.559	-1.892
P value	0.18	0.025	0.097	0.004	0.018	0.077
Cohen's d	0.541	0.934	-0.73	-1.373	-1.052	-0.724

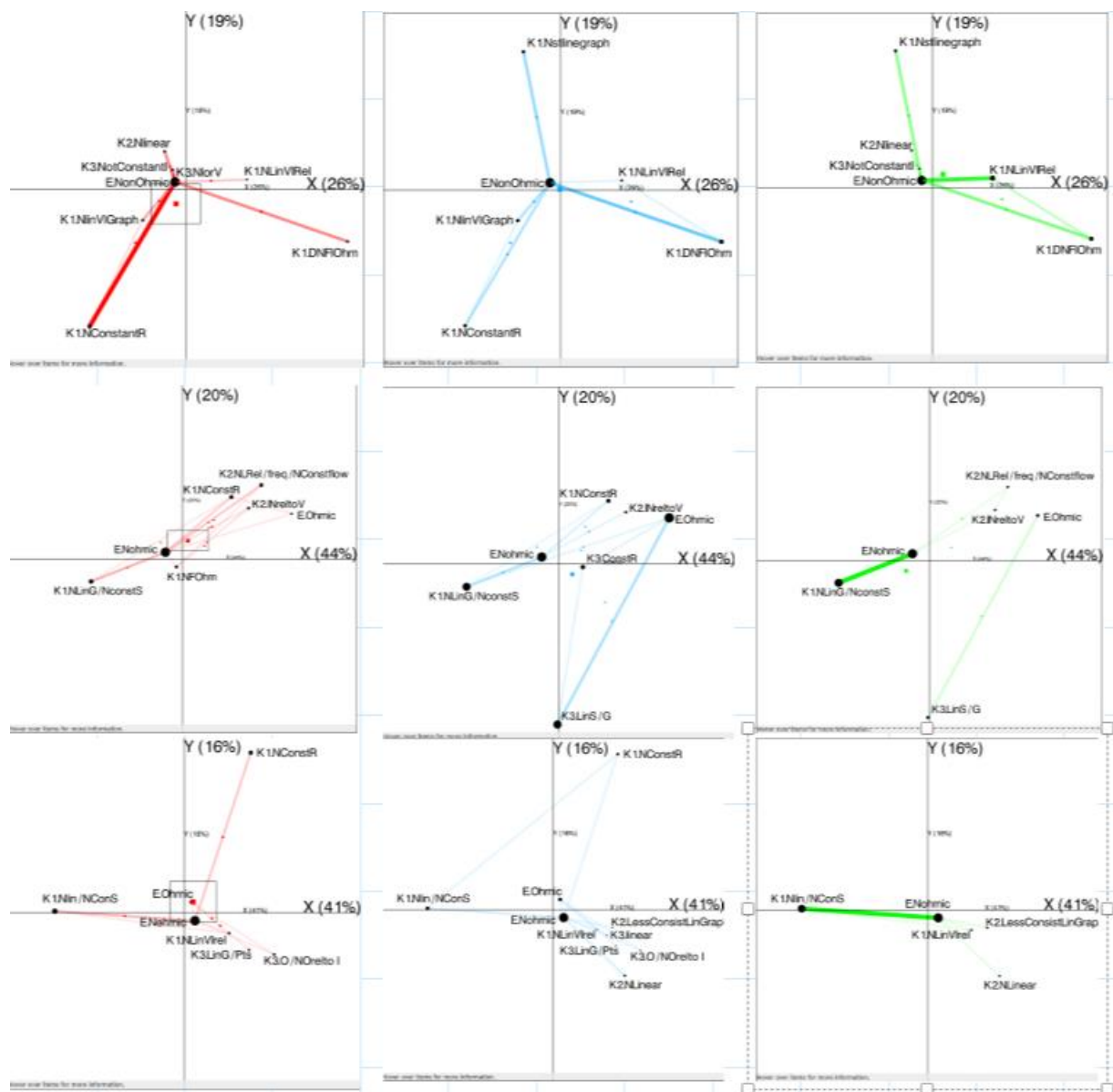


Figure A7. 3. Mean Network models of students by Low (red), Medium (blue), and High (green) Lab4 Quiz scores, for PreQ2 (upper 3graphs), PostQ2(center 3 graphs), and PostQ2 (lower 3 graphs).

Table A7. 4 T-test for Significance of Differences in Epistemic Networks of students by Low, Medium, and High Lab scores [PreQ2, Post Q1, and PostQ2]

	PreQ2		Post Q1		PostQ2	
	Dim 1	Dim 2	Dim1	Dim 2	Dim1	Dim2
Mean Low	0.193	-0.043	0.193	0.151	0.165	0.078
Mean Medium	0.015	0.061	0.026	-0.127	-0.138	-0.024
Mean High	-0.197	-0.03	-0.199	0.006	-0.138	-0.024
Medium-Low						
T statistic	-1.242	1.195	-0.874	-1.909	-1.635	-0.637
P value	0.228	0.243	0.391	0.07	0.116	0.536
Cohen's d	-0.498	0.461	-0.348	-0.788	-0.671	-0.299
High-Medium						
T statistic	-1.875	-0.653	-1.154	1.114	0.892	-0.159
P value	0.071	0.522	0.259	0.276	0.38	0.875
Cohen's d	-0.711	-0.258	-0.443	0.425	0.343	-0.062
High-Low						
T statistic	-2.657	0.098	-2.098	-1.096	-0.787	-0.697
P value	0.015	0.923	0.047	0.288	0.44	0.499
Cohen's d	-1.077	0.038	-0.864	-0.477	-0.33	-0.323



Figure A7. 4. Mean Network models of students by Low (red), Medium (blue), and High (green) Lab scores, for PreQ2 (upper 3 graphs), PostQ2 (center 3 graphs), and PostQ2 (lower 3 graphs).

Table A7. 5 T-test for Significance of Differences in Epistemic Networks of students by Low, Medium, and High Test1 scores [PreQ2, Post Q1, and PostQ2]

	PreQ2 Dim 1	Dim 2	Post Q1 Dim1	Dim 2	PostQ2 Dim1	Dim2
Mean Low	0.245	0.043	0.054	0.124	-0.025	-0.05
Mean Medium	-0.128	-0.056	-0.02	0.005	0.143	0.067
Mean High	-0.034	0.055	-0.038	-0.135	-0.224	-0.07
Medium-Low						
T statistic	-2.603	-0.823	-0.383	-0.977	0.858	0.859
P value	0.017	0.417	0.706	0.34	0.405	0.401
Cohen's d	-1.012	-0.274	-0.155	-0.388	0.383	0.345
High-Medium						
T statistic	0.867	1.102	-0.082	-0.988	-2.171	-1.668
P value	0.393	0.282	0.936	0.336	0.043	0.112
Cohen's d	0.299	0.324	-0.034	-0.41	-0.891	-0.498
High-Low						
T statistic	-2.017	0.143	-0.382	-1.677	-0.89	-0.186
P value	0.059	0.888	0.707	0.11	0.385	0.857
Cohen's d	-0.862	0.06	-0.174	-0.76	-0.412	-0.09

Appendix 8: T-tests and Network Graphs Experiment 4

Table A8. 1 T-test for Significance of Differences in Mean Epistemic Networks of students grouped by EBAPS scores, for PreQ, Post Q1, and PostQ2

	PreQ Dim 1	Dim 2	PostQ1 Dim1	Dim 2	PostQ2 Dim1	Dim2
Mean Low	0.044	0.062	-0.068	0.068	-0.098	-0.109
Mean Medium	0.04	0.068	-0.094	0.134	-0.109	-0.134
Mean High	0.451	-0.012	-0.089	0.128	-0.121	-0.235
Low-Medium						
T statistic	0.023	-0.194	1.978	-2.128	0.274	0.133
P value	0.982	0.849	0.068	0.047	0.79	0.897
Cohen's d	0.011	-0.088	0.876	-0.792	0.143	0.071
Medium-High						
0.258	0.451	-0.012	-0.317	0.142	0.359	0.642
0.798	-2.46	2.31	0.754	0.888	0.723	0.527
0.096	0.023	0.031	-0.127	0.057	0.145	0.259

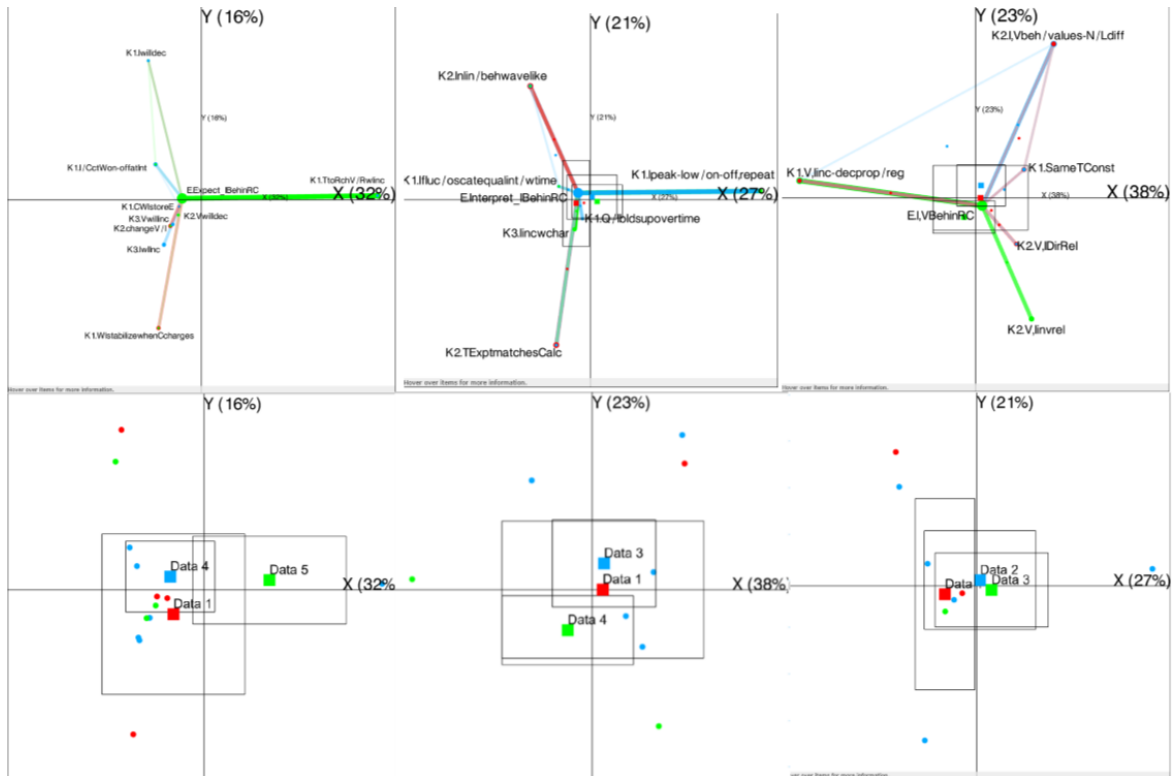


Figure A8. 1. Mean epistemic networks of students (upper 3 graphs) and centroid positions of all students (lower 3 graphs) sorted by EBAPS scores, as Low (red), Medium (blue), and High (green), for PreQ (left graphs), PostQ1(center graphs), and PostQ2 (right graphs).

Table A8. 2 T-test for Significance of Differences in Mean Epistemic Networks of students grouped by A3HWQuiz (Attempt1) scores, for PreQ, Post Q1, and PostQ2

	PreQ Dim 1	Dim 2	PostQ1 Dim1	Dim 2	PostQ2 Dim1	Dim2
Mean Low	-0.085	0.091	-0.07	0.068	-0.073	0.048
Mean Medium	0.381	0.001	-0.082	0.11	-0.133	-0.288
Mean High	0.082	0.059	-0.096	0.14	-0.105	-0.119
Low-Medium						
T statistic	-3.505	3.183	0.985	-1.384	2.11	2.69
P value	0.003	0.006	0.341	0.189	0.05	0.018
Cohen's d	-1.083	1.034	0.415	-0.523		
Medium-High						
0.258	1.863	-1.776	1.114	-0.822	-0.87	-1.103
0.798	0.074	0.088	0.277	0.419	0.393	0.281
0.096	0.69	-0.657	0.44	-0.32	-0.338	-0.429
Low-High						
T statistic	-1.851	1.577	2.495	-2.729	1.398	1.846
P value	0.082	0.13	0.025	0.013	0.179	0.086
Cohen's d	-0.508	0.48	0.917	-0.836	0.496	0.574

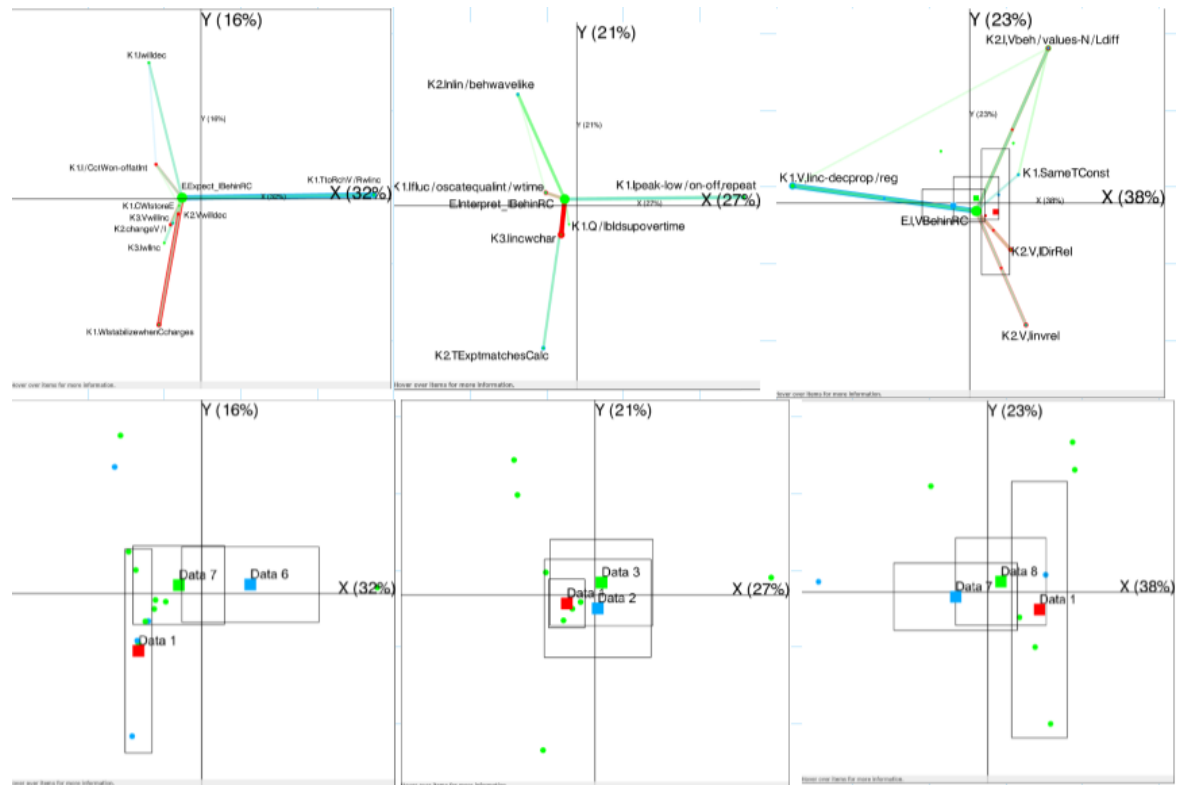


Figure A8. 2. Mean epistemic networks of students (upper 3 graphs) and centroid positions of all students (lower 3 graphs) sorted by A3Quiz scores, as Low (red), Medium (blue), and High (green), for PreQ (left graphs), PostQ1(center graphs), and PostQ2 (right graphs).

Table A8. 3 T-test for Significance of Differences in Mean Epistemic Networks of students grouped by Lab Quiz scores, for PreQ, Post Q1, and PostQ2

	PreQ		PostQ1		PostQ2	
	Dim 1	Dim 2	Dim1	Dim 2	Dim1	Dim2
Mean Low	-0.004	0.071	-0.085	0.105	-0.126	-0.195
Mean Medium	0.125	0.056	-0.083	0.122	-0.085	-0.068
Mean High	0.461	-0.017	-0.094	0.129	-0.132	-0.26
Low-Medium						
T statistic	-0.939	0.515	-0.105	-0.407	-1.372	-0.869
P value	0.358	0.612	0.918	0.688	0.184	0.395
Cohen's d	-0.37	0.203	-0.045	-0.175	-0.578	-0.368
Medium-High						
0.258	-1.738	1.873	0.704	-0.172	1.3	1.081
0.798	0.098	0.076	0.49	0.865	0.215	0.3
0.096	-0.748	0.803	0.304	-0.074	0.629	0.535
Low-High						
T statistic	-2.673	2.492	0.707	-0.664	0.152	0.327
P value	0.018	0.025	0.488	0.514	0.881	0.748
Cohen's d	-1.201	1.114	0.316	-0.297	0.072	0.154

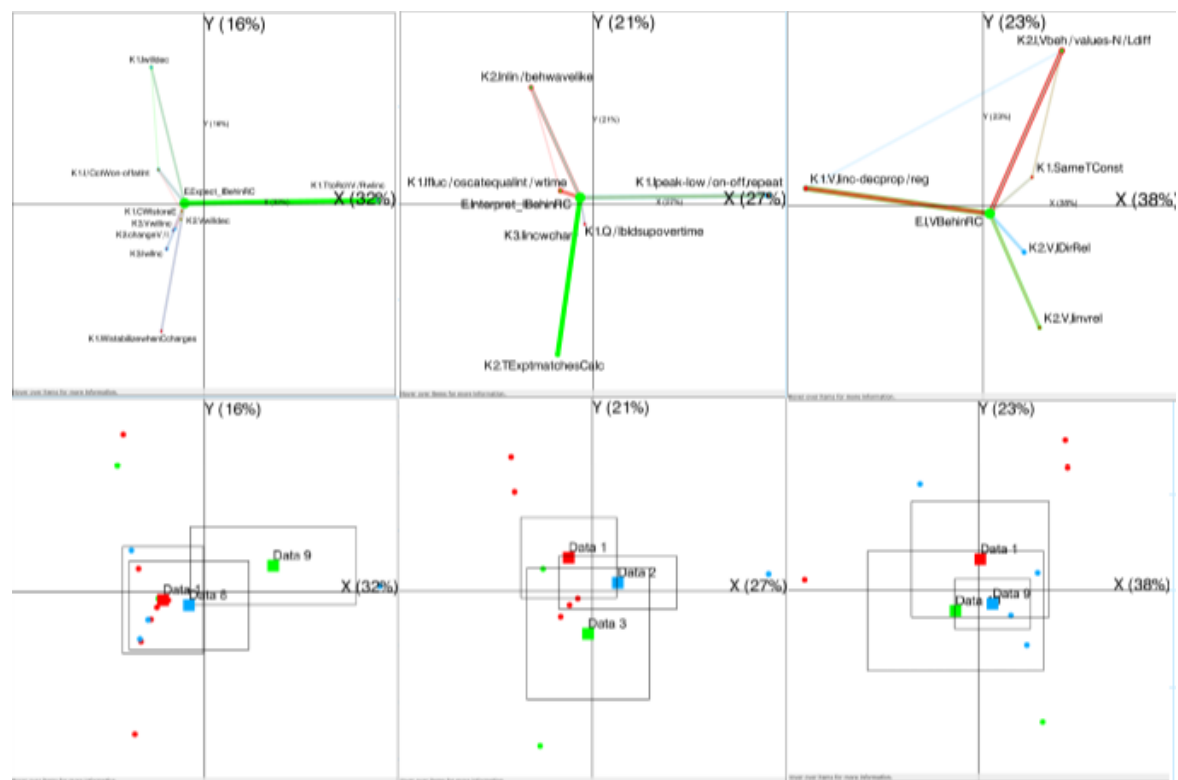


Figure A8. 3. Mean epistemic networks of students (upper 3 graphs) and centroid positions of all students (lower 3 graphs) sorted by Lab Quiz scores, as Low (red), Medium (blue), and High (green), for PreQ (left graphs), PostQ1(center graphs), and PostQ2 (right graph).

Table A8. 4 T-test for Significance of Differences in Mean Epistemic Networks of students grouped by Lab scores, for PreQ, Post Q1, and PostQ2

	PreQ Dim 1	Dim 2	PostQ1 Dim1	Dim 2	PostQ2 Dim1	Dim2
Mean Low	0.288	0.004	0.288	0.004	-0.112	-0.15
Mean Medium	0.176	0.049	0.08	0.063	-0.125	-0.214
Mean High	0.08	0.063	-0.101	0.162	-0.096	-0.118
Low-Medium						
T statistic	0.59	-1.233	1.116	-1.669	0.389	0.378
P value	0.562	0.23	0.278	0.11	0.702	0.709
Cohen's d	0.246	-0.50	0.493	-0.731	0.17	0.165
Medium-High						
0.258	0.588	-0.412	0.754	-0.997	-0.864	-0.588
0.798	0.562	0.684	0.46	0.332	0.397	0.563
0.096	0.232	-0.162	0.327	-0.436	-0.367	-0.25
Low-High						
T statistic	1.116	-1.669	2.61	-2.568	-0.447	-0.19
P value	0.278	0.11	0.019	0.024	0.66	0.852
Cohen's d	0.493	-0.731	1.173	-1.115	-0.206	-0.087

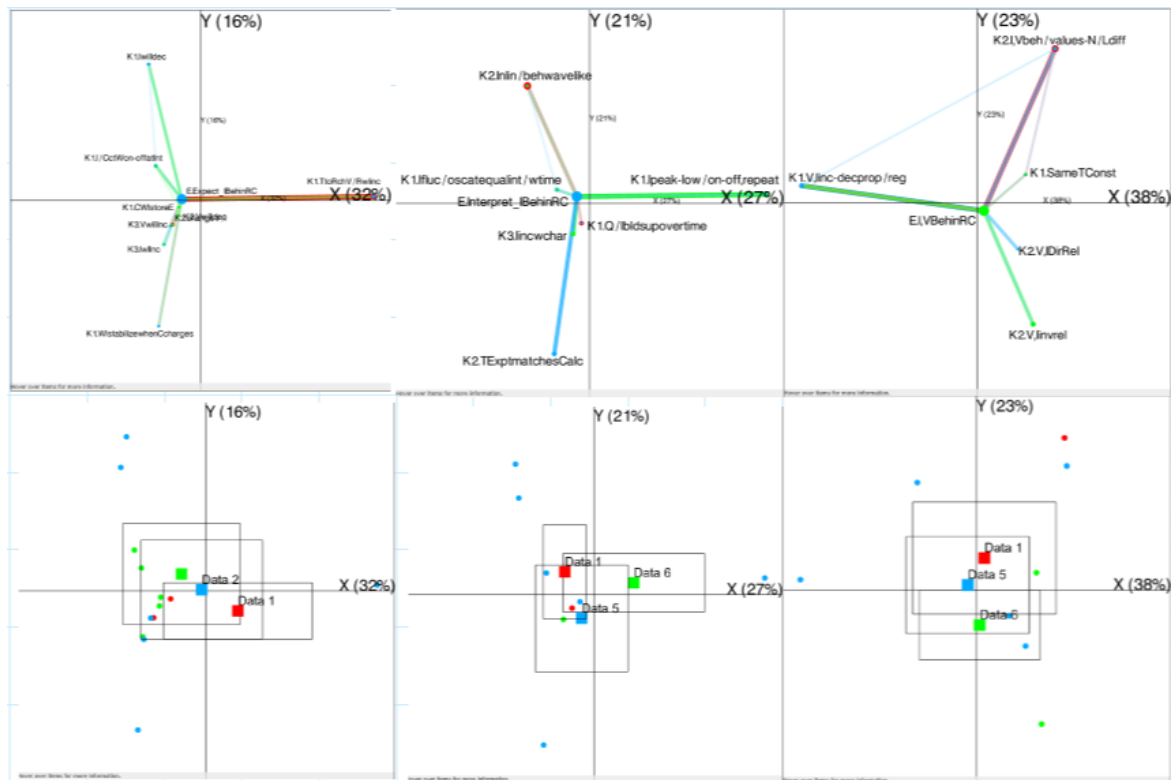


Figure A8. 4. Mean epistemic networks of students (upper 3 graphs) and centroid positions of all students (lower 3 graphs) sorted by Lab scores, as Low (red), Medium (blue), and High (green), for PreQ (left graphs), PostQ1(center graphs), and PostQ2 (right graphs).

Table A8. 5 T-test for Significance of Differences in Mean Epistemic Networks of students grouped by Test1 scores, for PreQ, Post Q1, and PostQ2

	PreQ Dim 1	Dim 2	PostQ1 Dim1	Dim 2	PostQ2 Dim1	Dim2
Mean Low	0.323	-0.001	-0.077	0.083	-0.144	-0.31
Mean Medium	0.071	0.065	-0.084	0.112	-0.076	0.011
Mean High	0.243	0.034	-0.104	0.164	-0.143	-0.326
Low-Medium						
T statistic	1.388	-1.934	0.581	-1.028	-1.902	-1.857
P value	0.186	0.071	0.568	0.315	0.092	0.102
Cohen's d	0.618	-0.817	0.238	-0.355	-1.093	-1.134
Medium-High						
0.258	-0.915	0.78	1.394	-1.192	2.034	2.143
0.798	0.378	0.45	0.184	0.253	0.066	0.058
0.096	-0.427	0.362	0.621	-0.537	1.008	1.102
Low-High						
T statistic	0.347	-0.778	1.794	-2.01	-0.016	0.069
P value	0.733	0.448	0.094	0.071	0.987	0.946
Cohen's d	0.168	-0.381	0.908	-0.992	-0.008	0.034

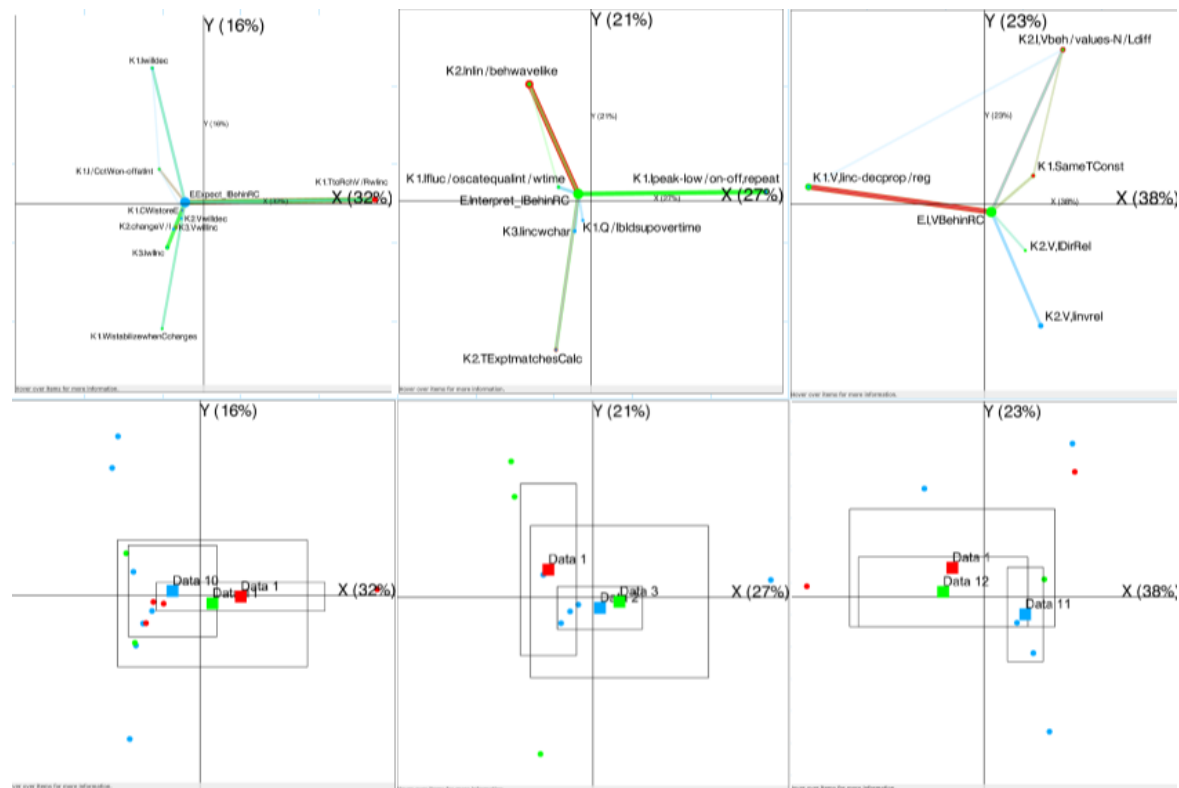


Figure A8. 5. Mean epistemic networks of students (upper 3 graphs) and centroid positions of all students (lower 3 graphs) sorted by Test1 scores, as Low (red), Medium (blue), and High (green), for PreQ (left graphs), PostQ1(center graphs), and PostQ2 (right graphs).

Appendix 9: R-Script

The following script was run on each of the 12 Rdata files exported to RStudio from ENAWebTool, to extract the required data columns listed in sub-section ‘Measures of Sophistication of Thinking’ on page 67.

```
library(dplyr)
getwd()
setwd("C:\\Users\\deepk\\Desktop\\RData")
newdata<-E1PreQID
df.name<-deparse(substitute(E1PreQID))
adkeyPre<-newdata[["adjacency.key"]]
mydataPre<-newdata[["set.information"]][["more.info"]][["outputs"]][["normed.unit.data"]]
mystackPre<-rbind(adkeyPre,mydataPre)
rlen<-nrow(mystackPre)
mystackPrenew<-mystackPre[c(1,2,5:rlen),]
mymetadataPre<-newdata[["set.information"]][["metadata"]]
myuserPre<-newdata[["set.information"]][["more.info"]][["data.used"]][["stanzasList"]]
datapositions<-newdata[["rotated.data"]]
combine1Pre<-cbind(myuserPre,datapositions)
combine2Pre<-cbind(combine1Pre,mydataPre)
combine3Pre<-cbind(mymetadataPre,combine2Pre)
write.csv(combine3Pre,paste(df.name, "_clean.csv", sep=""))
write.csv(mystackPrenew,paste(df.name, "_clean2.csv", sep=""))
write.csv(combine2Pre,paste(df.name, "_clean3.csv", sep=""))
```

Appendix 10: Excel Formulae to Compute S-scores

Excel function commands used to compute each of the seven scores of sophistication, from the data extracted from RStudio.

K1-centrality

=SUMIFS(AV4:BJ4,AV\$2:BJ\$2,"K1.*") +SUMIFS(AV4:BJ4,AV\$3:BJ\$3,"K1.*")

K2-centrality

=SUMIFS(AV4:BJ4,AV\$2:BJ\$2,"K2.*") +SUMIFS(AV4:BJ4,AV\$3:BJ\$3,"K2.*")

K3-centrality

=SUMIFS(AV4:BJ4,AV\$2:BJ\$2,"K3.*") +SUMIFS(AV4:BJ4,AV\$3:BJ\$3,"K3.*")

CCS

=SUMIFS(AV4:BJ4,AV\$2:BJ\$2,"E.*",AV\$3:BJ\$3,"K1.*")+SUMIFS(AV4:BJ4,AV\$2:BJ\$2,"K1*",AV\$3:BJ\$3,"K1.*")+SUMIFS(AV4:BJ4,AV\$2:BJ\$2,"K1*",AV\$3:BJ\$3,"E.*")

TCS

=SUMIFS(AV4:BJ4,AV\$2:BJ\$2,"K1.*",AV\$3:BJ\$3,"K2.*")+SUMIFS(AV4:BJ4,AV\$2:BJ\$2,"K1*",AV\$3:BJ\$3,"K3.*")+SUMIFS(AV4:BJ4,AV\$3:BJ\$3,"K1.*",AV\$2:BJ\$2,"K2*")+SUMIFS(AV4:BJ4,AV\$3:BJ\$3,"K1.*",AV\$2:BJ\$2,"K3*")

I1CS

=SUMIFS(AV4:BJ4,AV\$2:BJ\$2,"E.*",AV\$3:BJ\$3,"K2.*")+SUMIFS(AV4:BJ4,AV\$2:BJ\$2,"K2*",AV\$3:BJ\$3,"K2.*")+SUMIFS(AV4:BJ4,AV\$3:BJ\$3,"E.*",AV\$2:BJ\$2,"K2*")+SUMIFS(AV4:BJ4,AV\$3:BJ\$3,"K2.*",AV\$2:BJ\$2,"K2*")

I2CS

=SUMIFS(AV4:BJ4,AV\$2:BJ\$2,"E.*",AV\$3:BJ\$3,"K3.*")+SUMIFS(AV4:BJ4,AV\$2:BJ\$2,"K2*",AV\$3:BJ\$3,"K3.*")+SUMIFS(AV4:BJ4,AV\$3:BJ\$3,"E.*",AV\$2:BJ\$2,"K3*")+SUMIFS(AV4:BJ4,AV\$3:BJ\$3,"K2.*",AV\$2:BJ\$2,"K3*")+SUMIFS(AV4:BJ4,AV\$2:BJ\$2,"K3.*",AV\$3:BJ\$3,"K3.*")+SUMIFS(AV4:BJ4,AV\$3:BJ\$3,"K3.*",AV\$2:BJ\$2,"K3*")

Appendix 11: Levels of Sophistication Assignment Criteria

Table A11. 1 Connection strength distribution and levels of Sophistication: Experiment 1(PreQ and PostQ)

S - Level	Network Connection Strength and Co-occurrences in data	
	PreQ	Post Q
1	I2CS = 1	I2CS = 1
1	I1CS < I2CS <=1.15	
2	I1CS = 1	I1CS = 1
2	I1CS > 2	
3	CCS = TCS = I1CS < 1	CCS = TCS = I1CS < 1
4	CCS = 1	CCS = 1
5		CCS > 1

Table A11. 2 Connection strength distribution and levels of Sophistication: Experiment 2(PreQ, PostQ1, and PostQ2)

S - Level	Network Connection Strength and Co-occurrences in data		
	PreQ	PostQ1	PostQ2
1	I2CS = 1		I2CS = 1
1			I1CS < I2CS <=1.15
2	I1CS = 1	I1CS = 1	I1CS = 1
2	0.5 > CCS < TCS < I2CS < 2		CCS < TCS < I2CS < 2
		CCS < TCS < I1CS < 2	
		I1CS > 2	
2	CCS = TCS = I2CS < 1	CCS = TCS = I2CS < 1	CCS = TCS = I2CS < 1
2	CCS = TCS = I1CS < 1	CCS = TCS = I1CS < 1	CCS = TCS = I1CS < 1
3	CCS = 1	CCS = 1	CCS = 1
3	1 < CCS > TCS > I2CS		
			1.2 < CCS > TCS > I1CS < 0.5
3	CCS > 1	CCS > 1	CCS > 1
4			CCS > 3

Table A11. 3 Connection strength distribution and levels of Sophistication: Experiment 3(PreQ2, PostQ1, and PostQ2)

S - Level	Network Connection Strength and Co-occurrences in data		
	PreQ2	PostQ1	PostQ2
1	I2CS = 1	I2CS = 1	I2CS <= 1.15
		I2CS > 2	
	I1CS = 1	I1CS = 1	I1CS = 1
		I1CS > 2	
2		TCS = 1	
2		CCS = TCS = I1CS < 1	
2	CCS = 1	CCS = 1	CCS = 1
3	CCS > 1	CCS > 1	CCS > 1

Table A11. 4 Connection strength distribution and levels of Sophistication: Experiment 4(PreQ, PostQ1, and PostQ2)

S - Level	Network Connection Strength and Co-occurrences in data		
	PreQ	PostQ1	PostQ2
1	I2CS = 1	I2CS = 1	
	I2CS > 2		
1	I1CS = 1	I1CS = 1	I1CS = 1
2		CCS = TCS = I1CS<1	CCS = TCS = I1CS<1
3	CCS = 1	CCS = 1	CCS = 1
3	CCS > 1		

Appendix 12: Descriptive Statistics (Measures of Sophistication)

Table A12. 1 Descriptive Statistics Measures of Sophistication: Experiment 1(PreQ and PostQ)

	Strength	CCS	TCS	I1CS	I2CS	K1C	K2C	K3C
PreQ								
Mean		0.3453	0.3190	0.9505	0.1134	0.6644	1.2999	0.1134
Median		0.5773	0.5773	0.5773	0.0000	1.1547	1.1547	0.0000
Percent	.0000	42.1	44.7	10.5	89.5	42.1	10.5	89.5
	.5774	55.3	55.3	60.5				
	1.0000	2.6		7.9	5.3	2.6	7.9	5.3
	1.1547				5.3	55.3	60.5	5.3
	2.3094			18.4			18.4	
	3.6742			2.6			2.6	
PostQ								
Mean		1.0800	0.0137	0.0376	0.0476	1.2588	0.0513	0.0476
Median		1.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000
Percent	.0000	11.9	97.6	95.2	95.2	11.9	95.2	95.2
	.5774	2.4	2.4	2.4				
	1.0000	57.1		2.4	4.8	57.1	2.4	4.8
	1.1547					2.4	2.4	
	1.7321	28.6		18.4				
	2.3094			2.6				

Table A12. 2 Descriptive Statistics Measures of Sophistication: Experiment 2(PreQ, PostQ1, and PostQ2)

	Strength	CCS	TCS	I1CS	I2CS	K1C	K2C	K3C
PreQ								
Mean		0.7655	0.2084	0.0864	0.2998	1.0562	0.1505	0.4440
Median		1.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000
Percent	.0000	20.00	66.67	86.67	62.22	20.00	86.67	62.22
	0.4082	2.22			4.44			
	.5774	26.67	26.67	11.11	15.56			
	0.8165		6.67	2.22				
	1.0000	35.56			15.56	35.56	2.22	15.56
	1.1547				5.3	26.67	11.11	15.56
	1.2247	4.44				2.22		4.44
	1.6330				2.22			
	1.7321	11.11						
	2.3094					11.11		
	2.4495					4.44		2.22
PostQ1								
Mean		0.5599	0.0903	0.5850	0.0282	0.7206	0.6471	0.0563
Median		0.57735	0	0.57735	0	1	1	0
Percent	.0000	46.3	85.4	46.3	95.1	46.3	46.3	95.1
	0.4082	2.4						
	.5774	12.2	12.2	7.3	4.9			
	0.8165		2.4					
	1.0000	26.8		39		26.8	39	
	1.1547					12.2	7.3	4.9
	1.2247					2.4		
	1.6330			2.4				
	1.7321	12.2						
	2.3094			4.9		12.2	4.9	
	2.4495						2.4	
PostQ2								
Mean		1.1043	0.1307	0.1062	0.3932	1.5881	0.1715	0.4733
Median		1	0	0	0	1.154701	0	0
Percent	.0000	23.1	79.5	84.6	71.8	23.1	84.6	71.8
	0.4082	2.6		2.6				
	.5774	15.4	15.4	7.7	10.3			
	0.8165		5.1					
	1.0000	20.5		5.1	2.6	20.5	5.1	2.6
	1.1547				2.6	15.4	7.7	12.8
	1.2247	2.6				2.4	2.6	
	1.6330				2.6			
	1.7321	25.6						
	2.3094				10.3	25.6		10.3
	2.4495					2.6		2.6
	3.1623	10.3						
	5.0596					10.3		

Table A12. 3 Descriptive Statistics Measures of Sophistication: Experiment 3(PreQ2, PostQ1, and PostQ2)

	Strength	CCS	TCS	I1CS	I2CS	K1C	K2C	K3C
PreQ2								
Mean		0.6200	0.0000	0.0700	0.0900	0.6871	0.0700	0.0900
Median		1	0	0	0	1	0	0
Percent	.0000	46.5	100	93	90.7	46.5	93	90.7
	1.0000	41.9		7	9.3	41.9	7	9.3
	1.7321	11.6						
	2.3094					11.6		
PostQ1								
Mean		0.5549	0.0394	0.1722	0.2405	0.6232	0.2116	0.2405
Median		0.7887	0	0	0	1	0	0
Percent	.0000	47.5	95	85	82.5	45	82.5	82.5
	.5774	2.5	2.5	2.5				
	1.0000	42.5	2.5	10	12.5	45	12.5	12.5
	1.1547	2.5				5	2.5	
	1.7321	5						
	2.3094			2.5	5	5	2.5	5
PostQ2								
Mean		1.1043	0.1307	0.1062	0.3932	1.5881	0.1715	0.4733
Median		1.0000	0.0000	0.0000	0.0000	1.1547	0.0000	0.0000
Percent	.0000	23.1	79.5	84.6	71.8	23.1	84.6	71.8
	0.4082	2.6		2.6				
	.5774	15.4	15.4	7.7	10.3			
	0.8165		5.1					
	1.0000	20.5		5.1	2.6	20.5	5.1	2.6
	1.1547				2.6	15.4	7.7	12.8
	1.2247	2.6				2.6	2.6	
	1.6330				2.6			
	1.7321	25.6						
	2.3094				10.3	25.6		10.3
	2.4495					2.6		2.6
	3.1623	10.3						
	5.0596					10.3		

Table A12. 4 Descriptive Statistics Measures of Sophistication: Experiment 4(PreQ, PostQ1, and PostQ2)

	Strength	CCS	TCS	I1CS	I2CS	K1C	K2C	K3C
PreQ								
Mean		0.6701	0.0000	0.1300	0.1924	0.7005	0.1300	0.1924
Median		1	0	0	0	1	0	0
Percent	.0000	36.8	100	86.8	84.2	36.8	86.8	84.2
	1.0000	57.9		13.2	13.2	57.9	13.2	13.2
	1.7321	5.3						
	2.3094				2.6	5.3		2.6
PostQ1								
Mean		0.3405	0.0170	0.2523	0.0900	0.3575	0.2693	0.0900
Median		0.3405	0.0170	0.2523	0.0900	0	0	0
Percent	.0000	64.7	97.1	73.5	91.2	64.7	73.5	91.2
	.5774	2.9	2.9	2.9				
	1.0000	32.4		23.5	8.8	32.4	23.5	8.8
	1.1547					2.9	2.9	
PostQ2								
Mean		0.2862	0.0509	0.2862	0.0000	0.3372	0.3372	0.0000
Median		0	0	0	0	0	0	0
Percent	.0000	67.6	91.2	67.6	100	67.6	67.6	100
	0.4082							
	.5774	8.8	8.8	8.8				
	0.8165							
	1.0000	23.5		23.5		23.5	23.5	
	1.1547					8.8	8.8	

Appendix 13: Correlational Analyses Results

Table A13. 1 Spearman's rho Correlation Coefficient values between Sophistication Scores and Quantitative Scores

	CCS	TCS	I1CS	I2CS	K1C	K2C	K3C
PreQ							
EBAPS	0.293	0.169	-0.08	-0.089	0.209	0.044	-0.089
A3Attempt1	0.031	-0.047	-0.065	0.113	-0.026	-0.158	0.113
Quiz4	0.064	0.032	0.011	0.139	0.042	0.121	0.139
Lab4	0.081	-0.031	0.144	-0.079	0	0.142	-0.079
Test1	0.162	0.044	-0.003	0.015	0.079	0.058	0.015
Post Q							
EBAPS	0.073	-0.268	-.319*	0.096	-0.024	-.322*	0.096
A3Attempt1	0.200	0.147	0.210	-0.200	0.253	0.210	-0.200
Lab4	0.173	-0.201	-0.003	0.172	0.120	-0.016	0.172
Quiz4	0.227	0.137	-0.009	0.089	0.285	0.000	0.089
Test1	0.233	0.247	0.120	-0.061	.333*	0.131	-0.061

Table A13. 2 Spearman's rho Correlation Coefficient values between Sophistication Scores and Quantitative Scores

	K1C	K2C	K3C	CCS	TCS	I1CS	I2CS
PreQ							
EBAPS	-0.151	-0.074	-0.212	0.145	-0.245	-0.062	-0.186
A3Attempt1	0.043	0.087	-0.122	-0.006	-0.07	0.072	-0.063
Quiz4	-0.154	-0.187	0.107	-0.06	-0.129	-0.165	0.126
Lab4	-0.102	-0.039	-0.263	0.035	.300*	-0.047	-0.12
Test1	0.063	0.08	-0.165	0.031	-0.122	0.062	-0.096
PostQ1							
EBAPS	-0.043	0.099	0.01	-0.036	0.006	0.073	0.01
A3Attempt1	0.063	0.279	-0.083	-0.031	0.207	0.181	-0.083
Quiz4	0.212	-0.177	-0.005	0.24	-0.024	-0.185	-0.005
Lab4	0.217	-0.032	-0.019	0.155	0.108	-0.106	-0.019
Test1	0.151	0.182	0.203	-0.035	.421**	0.065	0.203
PostQ2							
EBAPS	0.154	0.097	0.167	0.009	0.201	0.08	0.142
A3Attempt1	0.282	-0.141	-0.208	.317*	-0.034	-0.164	-0.203
Lab4	-0.214	.407*	-0.06	-0.213	-0.046	.416**	-0.024
Quiz4	0.046	-0.037	-0.146	0.076	-0.027	-0.026	-0.168
Test1	-0.028	0.021	0.019	0.008	0.048	-0.04	0.032

Table A13. 3 Spearman's rho Correlation Coefficient values between Sophistication Scores and Quantitative Scores

	K1C	K2C	K3C	CCS	TCS	I1CS	I2CS
PreQ2							
EBAPS	0.198	-0.019	-0.053	0.198	.	-0.019	-0.053
A3Quiz	-0.037	-0.109	0.141	-0.037	.	-0.109	0.141
Quiz4	0.011	-0.104	-0.242	0.011	.	-0.104	-0.242
Lab4	0.139	-0.267	-0.189	0.139	.	-0.267	-0.189
Test1	-0.233	0.067	-0.065	-0.233	.	0.067	-0.065
PostQ1							
EBAPS	-0.032	-0.149	0.239	-0.079	0.173	-0.17	0.239
A3Attempt1	-0.212	-0.136	0.105	-0.081	.354*	-0.005	0.105
Quiz4	0.059	-.383*	0.211	0.174	-0.255	-.351*	0.211
Lab4	0.12	-.335*	-0.148	0.185	-0.14	-.352*	-0.148
Test1	0.021	-0.11	0.107	0.063	-0.102	-0.123	0.107
PostQ2							
EBAPS	-0.285	.351*	0.089	-0.285	.	.351*	0.089
A3Attempt1	-0.07	0.007	0.186	-0.07	.	0.007	0.186
Quiz4	0.094	0.146	-0.116	0.094	.	0.146	-0.116
Lab4	-0.165	-0.155	0.19	-0.165	.	-0.155	0.19
Test1	0.147	-0.179	0.07	0.147	.	-0.179	0.07

Table A13. 4 4 Spearman's rho Correlation Coefficient values between Sophistication Scores and Quantitative Scores

	K1C	K2C	K3C	CCS	TCS	I1CS	I2CS
PreQ							
overall	0.237	0.011	-0.126	0.237	.	0.011	-0.126
A3Attempt1	-0.094	-0.011	0.156	-0.094	.	-0.011	0.156
Quiz4	0.266	0.004	-0.261	0.266	.	0.004	-0.261
Lab4	0.318	0.097	-0.169	0.318	.	0.097	-0.169
Test1	0.115	-0.122	0.113	0.115	.	-0.122	0.113
PostQ1							
overall	0.169	-0.232	0.321	0.221	-0.167	-0.21	0.321
A3Attempt1	0.277	0.165	-0.216	0.253	0.163	0.141	-0.216
Quiz4	-0.077	0.138	0.022	-0.024	-0.226	0.193	0.022
Lab4	0.27	-0.109	.352*	0.277	0.036	-0.123	.352*
Test1	0.264	-0.003	0.016	0.221	0.233	-0.052	0.016
PostQ2							
overall	-0.001	-0.122	.	0.067	-0.155	-0.083	.
A3Attempt1	0.114	0.047	.	0	0.32	-0.08	.
Quiz4	0.051	-0.106	.	0.077	-0.038	-0.111	.
Lab4	0.774	0.549	.	0.665	0.833	0.531	.
Test1	0.113	0.061	.	0.117	0.043	0.054	.

Appendix 14: Cronbach's alpha Results

Table A14. 1 Inter-item Correlation and Cronbach's Alpha between (CCS, I1CS, TCS); (I1CS, K2C); (I2CS, K3C): Experiment 1

	CCS	TCS	K1C	Cronbach's Alpha	I1CS	K2C	I2CS	K3C	Cronbach's Alpha
PreQ									
CCS	1.000	0.854	0.965		I1CS	1.000	0.946		
TCS	0.854	1.000	0.961		K2C	0.946	1.000		0.967
K1C	0.965	0.961	1.000	0.923	I2CS		1.000	1.000	
					K3C		1.000	1.000	1.000
PostQ1									
CCS	1.000	-	0.968		I1CS	1.000	0.942		
		0.151							
TCS	-	1.000	-0.022		K2C	0.942	1.000		0.951
	0.151								
K1C	0.968	-	1.000	0.704	I2CS		1.000	1.000	
		0.022							
					K3C		1.000	1.000	1.000

Table A14. 2 Inter-item Correlation and Cronbach's Alpha between (CCS, I1CS, TCS); (I1CS, K2C); (I2CS, K3C): Experiment 2

	CCS	TCS	K1C	Cronbach's Alpha		I1CS	K2C	I2CS	K3C	Cronbach's Alpha
PreQ										
CCS	1.000	-	0.877		I1CS	1.000	0.952			
		0.122								
TCS	-	1.000	0.328		K2C	0.952	1.000			0.910
	0.122									
K1C	0.877	0.328	1.000	0.696	I2CS			1.000	0.914	
					K3C			0.914	1.000	0.926
PostQ1										
CCS	1.000	-	0.954		I1CS	1.000	0.962			
		0.013								
TCS	-	1.000	0.238		K2C	0.962	1.000			0.889
	0.013									
K1C	0.954	0.238	1.000	0.734	I2CS			1.000	1.000	
					K3C			1.000	1.000	0.978
PostQ2										
CCS	1.000	-	0.968		I1CS	1.000	0.942			
		0.151								
TCS	-	1.000	-0.022		K2C	0.942	1.000			0.951
	0.151									
K1C	0.968	-	1.000	0.704	I2CS			1.000	1.000	
		0.022								
					K3C			1.000	1.000	1.000

Table A14. 3 Inter-item Correlation and Cronbach's Alpha between (CCS, I1CS, TCS); (I1CS, K2C); (I2CS, K3C): Experiment 3

	CCS	TCS	K1C	Cronbach's Alpha	I1CS	K2C	I2CS	K3C	Cronbach's Alpha
PreQ2									
CCS	1.000		0.982		I1CS	1.000	1.000		
TCS					K2C	1.000	1.000		1.000
K1C	0.982		1.000	0.982	I2CS		1.000	1.000	
					K3C		1.000	1.000	1.000
PostQ1									
CCS	1.000	- 0.137	0.943		I1CS	1.000	0.933		
TCS	- 0.137	1.000	0.153		K2C	0.933	1.000		0.964
K1C	0.943	0.153	1.000	0.713	I2CS		1.000	1.000	
					K3C		1.000	1.000	1.000
PostQ2									
CCS	1.000	0.983			I1CS	1.000	1.000		
TCS					K2C	1.000	1.000		1.000
K1C	0.983	1.000		0.983	I2CS		1.000	1.000	
					K3C		1.000	1.000	1.000

Table A14. 4 Inter-item Correlation and Cronbach's Alpha between (CCS, I1CS, TCS); (I1CS, K2C); (I2CS, K3C): Experiment 4

	CCS	TCS	K1C	Cronbach's Alpha	I1CS	K2C	I2CS	K3C	Cronbach's Alpha
PreQ									
CCS	1.000		0.982		I1CS	1.000	1.000		
TCS					K2C	1.000	1.000		1.000
K1C	0.982		1.000	0.987	I2CS		1.000	1.000	
					K3C		1.000	1.000	1.000
PostQ1									
CCS	1.000	0.088	0.980		I1CS	1.000	0.977		
TCS	0.088	1.000	0.286		K2C	0.977	1.000		0.987
K1C	0.980	0.286	1.000	0.763	I2CS		1.000	1.000	
					K3C		1.000	1.000	1.000
PostQ2									
CCS	1.000	0.212	0.945		I1CS	1.000	0.945		
TCS	0.212	1.000	0.520		K2C	0.945	1.000		0.967
K1C	0.945	0.520	1.000	0.796	I2CS				
					K3C				

Appendix 15: Inter-Rater Agreement & Code List

Table A15. 1Inter-rater agreement

	Experiment 1		Experiment 2		Experiment 3		Experiment 4	
	Rater1	Rater 2	Rater1	Rater 2	Rater1	Rater 2	Rater1	Rater 2
PreQ								
1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	0.5	1
3	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	0.5
5	1	1	1	1	1	1	1	1
PostQ1								
6	0.67	0.67	1	1	0.5	1	0.67	1
7	1	1	1	1	1	1	1	1
8	1	1	1	1	0.67	1	1	1
9	1	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1
PostQ2								
11			0.67	1	1	0.5	1	1
12			1	1	1	1	1	1
13			1	1	1	1	1	1
14			1	0.5	1	1	0.5	1
15			1	1	1	1	1	1
Average	0.967	0.967	0.98	0.97	0.94	0.97	0.91	0.97

Table A15. 2 List of Codes by Experiment

Experiment 1	Experiment 2	Experiment 3	Experiment 4
K2.Shock/IncS/ Elcu/Dang	E.ChangeC	K2.ActasRinCct	E.IBehinRC
K2.V/I/Qwlinc	K1.ChangeD	K1.Const R	K1.Iwilldec
K3.V/I/Qdec	K2.ChangeNP	K3.Const I	K3.Vwillinc
K2.WConductor	K2.ChangeOA	k1.FLOhmL	K1.Wlstabilizewh enCharges
K2.WNConduct or	K1.ChangeA	K2.Linear	K2.changeV/I
K1.wetnessdecsk in/BR	K3.ChangeCct/Plsi/sh/ori/W	K1.LinGraph/S	K1.TtoRchV/Rwl inc
K1.WHBetterco nductor	K3.ChangeI/V/R/Q	K1.LinVIRel	K3.IwlInc
K1.WetnessdecR	K2.sepFoilsbypgs	E.OhmicMat	K1.I/CctWon- offatInt
K1.WetnessincC ond/I	E.InterpretActivity_RelC,A,D	K3.NlorV	K1.CWlstoreE
K1.dryHHhigher R	K1.CdecwA	K1.DNFLOhm	K2.Vwilldec
K1.I,RvarywWet ness	K1.CdecwIncD	K1.NLinVIRel	K1.Ipeak-low/on- off,repeat
K2.Noshock	K3.CN/lchangewA	K1.NConstantR	K1.Q/Ibldsupover time
K3.Wetness	K1.CIncwA	K1.NlinGraph/N constS	K1.Ifluc/oscatequ alint/wtime
INcR,decCond	K1.CincwDecD	E.NonOhmic	K2.TExptmatches Calc
E.Interpret_Rel	K3.R/VincwIncA	K3.NotConstantI	K2.Inlin/behwave like
Wetness,R	K3.R/V/WdecwIncD	K2.Nlinear	K3.Iincwchar
	K1.C,dinvrel	K2.NLRel/freq/ NConstflow	E.I,VBehinRC
	K1.C,AdirRel	K2.INreltoV	K2.V,IDirRel
	K2.CdecwIncNP	K2.LessConsist LinGraph	K2.I,Vbeh/values -N/Ldiff
	K3.CincwdecOA	K3.LinG/Pts	K1.SameTConst
	K3.CdecwIncOA	K3.linear	K1.V,Iinc- decprop/reg
		K3.O/NOrelto I	K2.V,Iinvrel
		K3.ConstR	
		K3.LinS/G	

VITA

Deepa Phadnis-Deshpande was born in Maharashtra, India, to Prabhakar and Snehal Phadnis. She is the first of three children: Milind and Nitin. She completed her schooling from Kendriya Vidyalaya Southern Command, in Pune, India, and went on to graduate with a Master's in Physics, from the University of Pune, in 1981. After completing her graduate studies in Physics, with a specialization in Electronics, she worked as Production and Quality-in-charge at Servoelec Industries, in Pune, for three years. Her marriage to Rajiv, and the birth of her first child, Ruchi, took her away from her career, for five years. She resumed her career as a Post-Graduate Teacher of Physics at the Junior College of IIT Powai, in Mumbai, India, in the year 1991. During her tenure from 1991 to 2007, she gained extensive experience of teaching physics, working as Board Examiner, and managing several school committees, to implement projects such as laboratory upgradation, scholarship institution, and technology integration. She was in-charge of the Physics Laboratory, including its inventory and resource planning. She was actively involved in mentoring students to prepare them for participating in Science Fairs, and Physics Olympiads. When her spouse Rajiv was transferred to the US by his company, she chose to retire early from her job in 2007.

After coming to the US, she began taking online courses in Education technology from the Pennsylvania State University. The field of study caught her interest, and she enrolled in the online master's program in Education Technology, offered by Penn State. She graduated in 2012, and then went on to join University of Tennessee's Doctoral program in Learning Environments and Education Studies, in 2013. She successfully defended her doctoral thesis in July 2018. Her research interests fall at the intersection of physics education and instructional technology. For her master's thesis, she examined the effect of interacting with simulation, and modelling with spreadsheet, on high school students' ability to solve complex problems in projectile motion. For her doctoral thesis, she has developed and successfully tested a technology-enabled assessment framework to assess the sophistication of learners' thinking in studio physics. As part of her assistantships, she taught an undergraduate course on 'international education' for three years, and assisted faculty with research and teaching. She is currently assisting a math education faculty member, with analysis of video data, to understand the effect of complex instruction and 'teaching math for social justice' on learners' equity of access to math. Her future plans include looking for post-doctoral opportunities in physics education research and faculty positions in STEM Education and/Education Technology.