UNIVERSITY OF MACEDONIA
DEPARTMENT OF APPLIED INFORMATICS
POSTGRADUATE PROGRAMME IN ARTIFICIAL INTELLIGENCE AND DATA ANALYTICS

EPISTEMIC NETWORK ANALYSIS: LITERATURE REVIEW AND CASE STUDY

A dissertation
by

Aliki Christou

Thessaloniki, September 2021

EPISTEMIC NETWORK ANALYSIS: LITERATURE REVIEW AND CASE STUDY

Aliki Christou

Bachelor of Applied Informatics, University of Macedonia, 2019

Dissertation

Submitted in partial fulfillment of the requirements for
THE POSTGRADUATE PROGRAMME IN ARTIFICIAL INTELLIGENCE AND DATA ANALYTICS

Supervisor Professor
Maya Satratzemi

Approved by the three-member Examination Committee on 24/09/2021

Maya Satratzemi	Georgia Koloniari	Panagiotis Papadimitriou
Professor	Assistant Professor	Assistant Professor

Aliki Christou

....

Abstract

Computer programming is a creative but complex task and findings have shown that it can be facilitated with collaboration. Advances in network and collaboration technologies allow the development of powerful collaborative tools, while communication is considered as one of the main factors for successful collaboration, and as such, chats provide rich information on the process of collaboration. This Thesis consists of two parts. The first one is a literature review for Epistemic Network Analysis (ENA) to analyze this method and also investigate the applications of it in different fields. The second part is two case studies for which we used ENA to analyze chat to see if we can detect differences between the connections made by Computer Science undergraduate students with different performance levels in an Object-Oriented Programming (OOP) course and their scores in a collaborative solving OOP assignment. The contribution of this Thesis is that we summarize and present the applications of ENA based on their characteristics and also propose a coding scheme of OOP elements using Epistemic Frame Theory in order to analyze how students are collaborated using chat messages for solving an OOP assignment and thus shade light on what type of connections are made in the groups of students with different computer programming skills. The results were mixed concerning the significant differences between the collaborative discourse networks of groups with different programming skills but we drew some interesting findings on the characteristics of the epistemic networks that groups of students with different programming skills form. Finally, limitations and future research are discussed.

Keywords: Epistemic Network Analysis, chat, Literature Review, Collaborative Learning, Object-Oriented Programming, Learning Analytics.

Foreword – Special thanks

Throughout the writing of this dissertation I have received a great deal of support and assistance. I would first like to thank my supervisor, Professor Maya Satratzemi for her invaluable advice, continuous support, and patience during my thesis. I would also like to thank Professor Stelios Xinogalos for his contribution during the research part, as a supervisor for some of the Zoom Meeting rooms. Also, I extend my thanks to all my colleagues for their continuous encouragement and support. In addition, I would like to thank my parents and my brother for their tremendous support and hope they had given to me. You are always there for me. Finally, I could not have completed this dissertation without the support of my friends who provided stimulating discussions as well as happy distractions to rest my mind outside of my research.

Table of contents

1 Introduction 1

1.1 Problem – Importance of the topic 1

1.2 Aim – Objectives 2

1.3 Research Questions 2

1.4 Contribution 3

1.5 Basic terminology 3

1.6 Structure of the Thesis 4

2 Epistemic Network Analysis (ENA) 5

2.1 Introduction 5

2.2 Cognitive Network Analysis 5

2.3 Main Characteristics of ENA 6

2.4 A Simple Comparison with ENA 7

Literature review – Theoretical background 12

2.5 Introduction 12

2.6 Learning Analysis for Education 13

2.6.1 Educational Collaborative Games 25

2.7 Teachers' Learning Analysis 27

2.8 Other Applications of ENA 31

2.9 Application of ENA for Medical Analysis 32

3 Case Study 35

3.1 Introduction 35

3.2 Methodology 35

3.2.1 Log Data 39

3.2.2 Coding scheme 40

3.2.3 Research Questions 41

3.3 Results 42

3.3.1 RQ1a. What types of connections between codes are made by each student's Group? 42

3.3.2 RQ1b. Is there a significant difference between the discourse networks of different students' Groups? 51

3.3.3 RQ2a. and RQ2b. 76

3.3.4 RQ3. Is there a significant difference between the discourse networks of groups of the same Category? 84
3.3.5 RQ4. Is there a significant difference between the discourse networks of the categories of the Groups based on the fundamental OOP concepts they used? 110
3.4 Summary 130
3.4.1 RQ1a. What types of connections between codes are made by each students' Group? 130
3.4.2 RQ1b. Is there a significant difference between the discourse networks of Group1 and the other nine different students' Groups? 132
3.4.3 RQ2a. What types of connections between codes are made by groups in the High-to-High category? What types of connections between codes are made by Groups in the High/Low-to-High category? What types of connections between codes are made by Groups in the High/Low-to-Low category? 132
3.4.4 RQ2b. Is there a significant difference between the discourse networks of groups of the three categories: High-to-High, High/Low-to-High and High/Low-to-Low? 133
3.4.5 RQ3. Is there a significant difference between the discourse networks of groups of the same Category? 134
3.4.6 RQ4. Is there a significant difference between the discourse networks of the categories of the Groups based on the fundamental OOP concepts they used? 135
4 Conclusion 137
4.1 Summary and conclusions 137
4.2 Research limitations 139
4.3 Future extensions 139
References 140
Annex A – Paper presentation at the ICERI2021 conference 145

List of Figures

Figure 2-1 The cognitive network of Student A 8
Figure 2-2 The cognitive network of Student B 9
Figure 2-3 The comparison network for Student A and Student B 9
Figure 2-4 The comparison of the students based on the positions of the centroids 11
Figure 3-1 Log Data 39
Figure 3-2 Network of the Group1 42
Figure 3-3 Network of the Group2 43
Figure 3-4 Network of the Group3 44
Figure 3-5 Network of the Group4 45
Figure 3-6 Network of the Group5 46
Figure 3-7 Network of the Group6 47
Figure 3-8 Network of the Group7 48
Figure 3-9 Network of the Group8 49
Figure 3-10 Network of the Group9 50
Figure 3-11 Network of the Group10 51
Figure 3-12 Centroids of the Ten Groups 52
Figure 3-13 The results of the Independent T-Test for the first dimension 53
Figure 3-14 The results of the Independent T-Test for the second dimension 53
Figure 3-15 Pearson's and Spearman's R for the Goodness of Fit 54
Figure 3-16 The Network of the centroid for Group1 54
Figure 3-17 Comparison of Group1 and Group2 54
Figure 3-18 The Network of the centroid for Group2 54
Figure 3-19 The results of the Independent T-Test for the first dimension 56
Figure 3-20 The results of the Independent T-Test for the second dimension 56
Figure 3-21 Pearson's and Spearman's R for the Goodness of Fit 56
Figure 3-22 The Network of the centroid for Group1 57
Figure 3-23 Comparison of Group1 and Group3 57
Figure 3-24 The Network of the centroid for Group3 57
Figure 3-25 The results of the Independent T-Test for the first dimension 58
Figure 3-26 The results of the Independent T-Test for the second dimension 58
Figure 3-27 Pearson's and Spearman's R for the Goodness of Fit 59

Figure 3-28 The Network of the centroid for Group1 59
Figure 3-29 Comparison of Group1 and Group4 59
Figure 3-30 The Network of the centroid for Group4 59
Figure 3-31 The results of the Independent T-Test for the first dimension 61
Figure 3-32 The results of the Independent T-Test for the second dimension................. 61
Figure 3-33 Pearson’s and Spearman’s R for the Goodness of Fit 61
Figure 3-34 The Network of the centroid for Group1 62
Figure 3-35 Comparison of Group1 and Group5 62
Figure 3-36 The Network of the centroid for Group5 62
Figure 3-37 The results of the Independent T-Test for the first dimension 63
Figure 3-38 The results of the Independent T-Test for the second dimension................. 63
Figure 3-39 Pearson’s and Spearman’s R for the Goodness of Fit 64
Figure 3-40 The Network of the centroid for Group1 64
Figure 3-41 Comparison of Group1 and Group6 64
Figure 3-42 The Network of the centroid for Group6 64
Figure 3-43 The results of the Independent T-Test for the first dimension 66
Figure 3-44 The results of the Independent T-Test for the second dimension................. 66
Figure 3-45 Pearson’s and Spearman’s R for the Goodness of Fit 66
Figure 3-46 The Network of the centroid for Group1 67
Figure 3-47 Comparison of Group1 and Group7 67
Figure 3-48 The Network of the centroid for Group7 67
Figure 3-49 The results of the Independent T-Test for the first dimension 68
Figure 3-50 The results of the Independent T-Test for the second dimension................. 68
Figure 3-51 Pearson’s and Spearman’s R for the Goodness of Fit 69
Figure 3-52 The Network of the centroid for Group1 69
Figure 3-53 Comparison of Group1 and Group8 69
Figure 3-54 The Network of the centroid for Group8 69
Figure 3-55 The results of the Independent T-Test for the first dimension 71
Figure 3-56 The results of the Independent T-Test for the second dimension................. 71
Figure 3-57 Pearson’s and Spearman’s R for the Goodness of Fit 71
Figure 3-58 The Network of the centroid for Group1 72
Figure 3-59 Comparison of Group1 and Group9 72
Figure 3-60 The Network of the centroid for Group9 72

Figure 3-61 The results of the Independent T-Test for the first dimension 73
Figure 3-62 The results of the Independent T-Test for the second dimension................. 73
Figure 3-63 Pearson's and Spearman's R for the Goodness of Fit 74
Figure 3-64 The Network of the centroid for Group1... 74
Figure 3-65 Comparison of Group1 and Group10 74
Figure 3-66 The Network of the centroid for Group10.. 74
Figure 3-67 The Centroids of the Three Categories Groups (Red: High-to-High, Purple: High/Low-to-High and Blue: High/Low-to-Low).. 76
Figure 3-68 Correlation for the 3 Categories 77
Figure 3-69 The results of the Independent T-Test for the first dimension 78
Figure 3-70 The results of the Independent T-Test for the second dimension................. 78
Figure 3-71 Network of groups in the High-to-High category.. 79
Figure 3-72 Comparison of High-to-High and High/Low-to-High group categories...... 79
Figure 3-73 Network of groups in the High/Low-to-High category 79
Figure 3-74 The results of the Independent T-Test for the first dimension 80
Figure 3-75 The results of the Independent T-Test for the second dimension................. 80
Figure 3-76 Network of groups in the High-to-High category.. 81
Figure 3-77 Comparison of High-to-High and High/Low-to-Low 81
Figure 3-78 Network of groups in the High/Low-to-Low category................................. 81
Figure 3-79 The results of the Independent T-Test for the first dimension 82
Figure 3-80 The results of the Independent T-Test for the second dimension................. 82
Figure 3-81 Network of groups in the High/Low-to-High category 83
Figure 3-82 Comparison of High/Low-to-High and High/Low-to-Low.......................... 83
Figure 3-83 Network of groups in the High/Low-to-Low category................................. 83
Figure 3-84 The Centroids networks of the Three Groups of the category High-to-High (RED: Group1, PINK: Group5, BLUE: Group8)... 84
Figure 3-85 The results of the Independent T-Test for the first dimension 86
Figure 3-86 The results of the Independent T-Test for the second dimension................. 86
Figure 3-87 Pearson's and Spearman's R for the Goodness of Fit 86
Figure 3-88 The Network of the centroid for Group1... 87
Figure 3-89 Comparison of Group1 and Group5 87
Figure 3-90 The Network of the centroid for Group5.. 87
Figure 3-91 The results of the Independent T-Test for the first dimension 88

Figure 3-92 The results of the Independent T-Test for the second dimension................. 88
Figure 3-93 Pearson's and Spearman's R for the Goodness of Fit 89
Figure 3-94 The Network of the centroid for Group1... 89
Figure 3-95 Comparison of Group1 and Group8 89
Figure 3-96 The Network of the centroid for Group8... 89
Figure 3-97 The results of the Independent T-Test for the first dimension 91
Figure 3-98 The results of the Independent T-Test for the second dimension................. 91
3-99 Pearson's and Spearman's R for the Goodness of Fit... 91
Figure 3-100 The Network of the centroid for Group5... 92
Figure 3-101 Comparison of Group5 and Group8 92
Figure 3-102 The Network of the centroid for Group8... 92
Figure 3-103 The Centroids of the Three Groups of the High/Low-to-High Category (GREEN: Group4, RED: Group7, PURPLE: Group9) 93
Figure 3-104 The results of the Independent T-Test for the first dimension 94
Figure 3-105 The results of the Independent T-Test for the second dimension............... 94
Figure 3-106 Pearson's and Spearman's R for the Goodness of Fit 95
Figure 3-107 The Network of the centroid for Group4... 95
Figure 3-108 Comparison of Group4 and Group7 95
Figure 3-109 The Network of the centroid for Group7... 95
Figure 3-110 The results of the Independent T-Test for the first dimension 97
Figure 3-111 The results of the Independent T-Test for the second dimension............... 97
3-112 Pearson's and Spearman's R for the Goodness of Fit... 97
Figure 3-113 The Network of the centroid for Group4... 98
Figure 3-114 Comparison of Group4 and Group9 98
Figure 3-115 The Network of the centroid for Group9... 98
Figure 3-116 The results of the Independent T-Test for the first dimension 99
Figure 3-117 The results of the Independent T-Test for the second dimension............... 99
Figure 3-118 Pearson's and Spearman's R for the Goodness of Fit 100
Figure 3-119 The Network of the centroid for Group7.. 100
Figure 3-120 Comparison of Group7 and Group9 100
Figure 3-121 The Network of the centroid for Group9.. 100
Figure 3-122 The Centroids of the Three Groups of High/Low-to-Low Category........ 102
Figure 3-123 The results of the Independent T-Test for the first dimension 103

Figure 3-124 The results of the Independent T-Test for the second dimension............. 103
3-125 Pearson's and Spearman's R for the Goodness of Fit... 104
Figure 3-126 The Network of the centroid for Group2... 104
Figure 3-127 Comparison of Group2 and Group3 104
Figure 3-128 The Network of the centroid for Group3... 104
Figure 3-129 The results of the Independent T-Test for the first dimension 106
Figure 3-130 The results of the Independent T-Test for the second dimension............. 106
Figure 3-131 Pearson's and Spearman's R for the Goodness of Fit 106
Figure 3-132 The Network of the centroid for Group2... 107
Figure 3-133 Comparison of Group2 and Group6 107
Figure 3-134 The Network of the centroid for Group6... 107
Figure 3-135 The results of the Independent T-Test for the first dimension 108
Figure 3-136 The results of the Independent T-Test for the second dimension............. 108
Figure 3-137 Pearson's and Spearman's R for the Goodness of Fit 109
Figure 3-138 The Network of the centroid for Group3... 109
Figure 3-139 Comparison of Group3 and Group6 109
Figure 3-140 The Network of the centroid for Group6... 109
Figure 3-141 The Centroids of the Four Group Categories (RED: Abstract, Inheritance, GREEN: Inheritance, PURPLE: Abstract, Inheritance, Comparator, BLUE: None) 111
Figure 3-142 The results of the Independent T-Test for the first dimension 112
Figure 3-143 The results of the Independent T-Test for the second dimension............. 112
Figure 3-144 Pearson's and Spearman's R for the Goodness of Fit 113
Figure 3-145 The Network of the centroid of groups used Abstract, Inheritance, Comparator 114
Figure 3-146 The Network of the centroid of groups used Abstract, Inheritance.......... 114
Figure 3-147 The results of the Independent T-Test for the first dimension 115
Figure 3-148 The results of the Independent T-Test for the second dimension............. 115
Figure 3-149 Pearson's and Spearman's R for the Goodness of Fit 116
Figure 3-150 The Network of the centroid for Inheritance 117
Figure 3-151 The Network of the centroid for Abstract, Inheritance............................. 117
Figure 3-152 The results of the Independent T-Test for the first dimension 118
Figure 3-153 The results of the Independent T-Test for the second dimension............. 118

Figure 3-154 Pearson's and Spearman's R for the Goodness of Fit 119
Figure 3-155 The Network of the centroid for None(-) 120
Figure 3-156 The Network of the centroid for Abstract, Inheritance............................ 120
Figure 3-157 The results of the Independent T-Test for the first dimension 122
Figure 3-158 The results of the Independent T-Test for the second dimension............. 122
Figure 3-159 Pearson's and Spearman's R for the Goodness of Fit 122
Figure 3-160 The Network of the centroid for Inheritance 123
Figure 3-161 The Network of the centroid for Abstract, Inheritance, Comparator 123
Figure 3-162 The results of the Independent T-Test for the first dimension 125
Figure 3-163 The results of the Independent T-Test for the second dimension............. 125
Figure 3-164 Pearson's and Spearman's R for the Goodness of Fit 125
Figure 3-165 The Network of the centroid for None 126
Figure 3-166 The Network of the centroid for Abstract, Inheritance, Comparator 126
Figure 3-167 The results of the Independent T-Test for the first dimension 128
Figure 3-168 The results of the Independent T-Test for the second dimension............. 128
Figure 3-169 Pearson's and Spearman's R for the Goodness of Fit 128
Figure 3-170 The Network of the centroid for None 129
Figure 3-171 The Network of the centroid for Inheritance 129

List of tables

Table 3-1: Applications of ENA in Learning Analysis for Education 14
Table 3-2: Applications of ENA in Teachers' Learning Analysis.................................... 27
Table 3-3: Applications of ENA in the field of Nutrition 31
Table 3-4: Applications of ENA in Medical Analysis 32
Table 4-1 Members and Groups Scores 37

Symbolisms

- ENA: Epistemic Network Analysis
- OOP: Object Oriented Programming

1 Introduction

1.1 Problem – Importance of the topic

Computer programming is a creative but complex task and findings have shown that it can be facilitated with collaboration (L. Silva et al, 2020). Many studies have been contucted concerning Pair Programming (PP) or team work (A. Cockburn et al, 2001), (L. Williams et al, 2003) and Distributed Pair Programming (DPP) (P. Baheti et al, 2002), (S.Xinogalos et al, 2019) in order to study how Computer Science/Software Engineering students benefit from working in pairs or teams, co-located in the case of PP or remotely in the case of DPP. All these studies highlighted that collaboration is feasible if an underlying infrastructure enables for all necessary interactions (T. Schümmer et al, 2009). Advances in network and collaboration technologies have allowed the development of powerful collaborative tools for supporting code development, and communication is considered as one of the main factors for successful collaboration. In the majority of these tools communication was possible via a textual chat. Most of the studies examined the method of communication: specific features of the tools, such as embedded chat, remote selection of code, gesturing features; communication strategies or styles; and students' satisfaction.

Eventhough, chats provide rich information on the process of collaboration all these studies concerning collaborative programming (PP or team Programming or DPP) have not analyzed chat data. Last decades researchers in the Learning Analytics field developed many tools to analyze chat data. Epistemic Network Analysis (ENA) (https://www.epistemicnetwork.org/) is a network analysis technique that analyses logfile data and other records of individual and collaborative learning. ENA offers powerful mechanisms to analyze collaboration discourse and links among relevant features of collaborative learning. ENA was originally developed to model theories of cognition, discourse, and culture which argue that the connections people make in discourse are a critical level of analysis (D.W. Shaffer et al, 2009). ENA models the connections between the discourse elements or codes by quantifying their co-occurrence producing a weighted network of co-occurrences. The frequency of the co-occurrence of two codes in a discourse is used to compute the strength of their association in a network (see section 2 for a more detailed description of ENA).

1.2 Aim – Objectives

The objectives of this Thesis are the following: (a) to summarize, present and discuss the applications of ENA with a Literature Review, (b) to conduct a case study in order to examine the types of connections between codes made by each student and groups of students during collaborative code development and (c) to compare either groups of students which are comprised by members of different levels of computer programming skills or groups with each other to find if there is significant difference between their discourse networks.

1.3 Research Questions

In order to investigate aims b) and c) we conducted a case study based on primary research. In this study we used ENA to analyze chat to see if we can detect differences between the connections made by students with different performance levels in an Object-Oriented Programming (OOP) course and their scores in a collaborative solving assignment. In order to examine the types of connections made between different types of groups based on their performance, we considered their mean grade in the course and their mean grade in the assignment. Students with a score above or below the mean course grade and groups with scores above or below the mean assignment grade were categorized as High or Low scoring, respectively. We had different categories of groups based on members' individual and group scores. There were groups of High student course score and High group assignment score (High-to-High category); groups of High or Low student course score and group High group assignment score (High/Low-to-High category); and groups of High or Low student course score and Low group assignment score (High/Low-to-Low category). Based on the above defined assumptions the Research Questions of this study are formulated as follows:

- RQ1a. What types of connections between codes are made by each students' Group?
- RQ1b. Is there a significant difference between the discourse networks of the different students' Groups?
- RQ2a. What types of connections between codes are made by groups in the High-to-High category? What types of connections between codes are made by Groups in the High/Low-to-High category? What types of connections between codes are made by Groups in the High/Low-to-Low category?

- RQ2b. Is there a significant difference between the discourse networks of groups of the three categories: High-to-High, High/Low-to-High and High/Low-to-Low?
- RQ3. Is there a significant difference between the discourse networks of groups of the same Category?
- RQ4. Is there a significant difference between the discourse networks of the categories of the Groups based on the fundamental OOP concepts they used?

1.4 Contribution

Although ENA has been used in chat analysis in many different disciplines no study has analyzed students' discourse when working collaboratively to solve an OOP assignment. In this study we propose a coding scheme of OOP elements using the Epistemic Frame Theory in order to analyze how students collaborate using chat messages to solve an OOP assignment, and in so doing, shed light on what type of connections are made in the groups of students, which are comprised of different levels of computer programming skills. This study's findings will contribute to further understanding of the developing and soft skills of the students through the collaborative code development.

1.5 Basic terminology

The basic terminology that is going to be used in this study is the following:

- Epistemic Network Analysis (ENA) is a method for identifying and quantifying connections among cognitive or other elements in coded data and displaying them in dynamic network models.
- Object-oriented programming (OOP) is a programming development methodology, supported by appropriate programming languages, where the handling of related data and the processes that affect them is done together, through a data structure that surrounds them as an autonomous entity with its own identity and characteristics. This data structure is called an object and is a real snapshot in the memory of a complex, and possibly user-defined, type of data called a class. The class specifies both data and the processes that affect them; this has been the primary innovation of CA.

1.6 Structure of the Thesis

This Thesis consists of five chapters. The first one is the introduction of the study. In the second one we present Epistemic Network Analysis (ENA) and its characteristics and the implementation of this method. In Chapter 3, a literature review on ENA applications is presented and discussed. The forth one presents the research conducted and a case study with two parts of it. The fifth and final chapter is the conclusions, the limitations and the future research of the study.

2 Epistemic Network Analysis (ENA)

2.1 Introduction

In this chapter, we explain how Epistemic Network Analysis (ENA) is implemented. Epistemic Network Analysis (ENA) is a technique that analyses logfile data and other records of individual and collaborative learning. This method is also used to quantify, visualize and interpret network data. It was originally developed to model theories of cognition, discourse, and culture, which argue that the connections people make in discourse are a critical level of analysis (D.W. Shaffer et al, 2009). ENA offers powerful mechanisms to analyze collaboration discourse and links among relevant features of collaborative learning. Specifically, it models the connections between the discourse elements or codes by quantifying their co-occurrence, thus producing a weighted network of co-occurrences. The frequency of the co-occurrence of two codes in a discourse is used to compute the strength of their connection in a network.

2.2 Cognitive Network Analysis

The models produced from ENA are visualized as networks (Shaffer, D. W. et al, 2016) (Shaffer, D. W., 2018(b)). Each node of the network is a code, while the network edges represent the co-occurrence of the connected codes. The connections can be quantified between the codes and that results in a weighted graph. When the edges that are darker and thicker indicate a stronger connection, meaning the study participant made those connections more often, and vice versa, edges that are lighter and thinner equal weaker connections as they were made less often. The final models produced by ENA enable the comparison of networks based on the nodes, the edges (connections), and the statistics which describe each network and its features.

ENA was originally developed to model cognitive networks based on a fundamental assumption of some theories of learning analytics: that the structure of connections between cognitive elements is more important than the mere presence or absence of these elements individually. Shaffer describes learning as the development of a scientific context: a pattern of correlations between knowledge, skills, mental habits, and other cognitive elements that characterize communities of practice or groups of people who share similar ways of shaping, exploring, and solving complex problems. He also emphasizes that learning is not defined by the possession of individual pieces of knowledge and other skills, but by the structure of the connections between them. Thus the basic assumption of ENA is that the structure of data links is more important in the analysis.

However, ENA can also model connection patterns in any system characterized by a complex network of dynamic relationships between a relatively small, fixed data set. Thus it is a method optimized for network analysis that is too large for parametric techniques with many variables, such as latent class models, but not so large that they require analysis only through summary statistics, as is the case with many traditional network analysis techniques. Subsequently, we will describe the theory behind ENA and the process by which ENA creates network models. The following drawbacks are identified for modeling these networks: the number of interactions increases exponentially concerning the number of nodes and the structure of the connections is a fundamental question about networks.

2.3 Main Characteristics of ENA

ENA is based on a distinct theory of learning: the epistemic frame hypothesis (D.W. Shaffer et al, 2004), (D.W. Shaffer et al, 2006). This theory suggests that any community of practice has a culture, and that this culture has a grammar, which is composed of:

- Skills: the things people within the community can do
- Knowledge: the perceptions people in the community have
- Identity: the way community members see themselves
- Values: the beliefs of the community members
- Epistemology: the credentials that explain actions or claims as valid within the community

This collection of skills, knowledge, identity, values, and epistemology forms the epistemic frame of the community. The Epistemic Frame Hypothesis claims that an epistemic frame binds together the grammar (i.e., skills, knowledge, values, identity, and epistemology) that one takes on as a member of a community of practice.

ENA has three basic assumptions: it is possible to identify distinct features in data, the data analyzed have a local structure and the network elements have a significant data feature on which they are connected. ENA has also the following three characteristics:

- Codes: the researched features of the data
- Units: can be either a group of participants or actions observed or a combination of the two, and
- Stanza: is part of the data in which the coexistence of the codes is examined.

To do this, ENA creates an adjacency matrix that depicts the co-occurrences of codes per stanza (Shaffer, D. W., 2018(a)). If a code co-occurs in a stanza, ENA assigns one, and zero if it does not. Then, the adjacency matrices are summed up into a cumulative adjacency matrix. Each cell of the final matrix displays the number of stanzas in which that unique pair of codes was observed. ENA then converts cumulative adjacency matrices into cumulative adjacency vectors by projecting them into a high-dimensional space. After that, it performs singular value decomposition on the normalized vectors, producing a rotation of the original high-dimensional space. This action provides a reduced number of dimensions that capture the maximum variance in the data.

There are two main techniques used for visualization in ENA spherical normalization and the dimensionality reduction (Shaffer, D. W.et al, 2016). Spherical normalization is accomplished by dividing each vector by its length. The resulting normalized vector quantifies for the unit the relative frequencies of co-occurences of codes independent of the number of stanzas in the model for any given unit. As for the second one in order to interpret and visualize the normalized adjacency vectors $\boldsymbol{N_u}$, ENA performs a dimensional reduction using Singular Value Decomposition. For each unit u in the data, ENA generates a point, $\boldsymbol{P_u}$, which is the position of the normalized vector $\boldsymbol{N_u}$, after decomposing a single value. To interpret the dimensions of this rotating space, ENA takes the codes from the original data - which corresponds to the network nodes - and uses an optimization routine to place them in the ENA space, so that for any u unit, its centroid model of the network corresponding to the cumulative neighborhood table C_u be as close as possible to the location of the point P_u.ENA calculates and reports the strength of the correlation between the centroids and the projected points in the model using both Pearson's r and Spearman's r.

2.4 A Simple Comparison with ENA

To understand ENA it would be helpful to provide an example. Consider Figure 2-1, which shows the cumulative cognitive network on an undergraduate student of the

Department of Applied Informatics (Student A), who participated in the research. The student’s network models the structure of connections between Skills, Knowledge, Values, Identities and Epistemology codes. This model of Student A shows strong connection between the skill of collaboration and the skill of design, the identity of supportive and the skill of collaboration, the identity of expert and the skill of collaboration, the skill of collaboration and the skill of data and the identity of supportive and the skill of design. The nodes in the networks of both students appear exactly at the same points in the view point for all students. Placing nodes in fixed positions allows the connection patterns to be compared on two or more networks, but also allows the interpretation of the projection space itself. Also, the nodes of the cognitive network that feature more connections have a larger size in the graph than those with fewer connections(Shaffer, D. W., 2016). The network is weighted which means that the darker and thicker lines indicate connections that this student made more often and the lighter and thinner lines indicate connections that this student made more rarely.

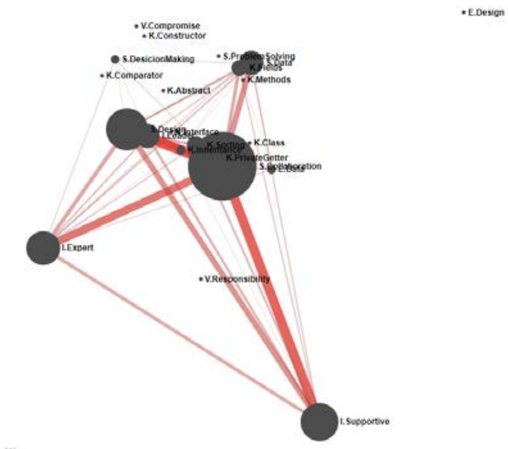


Figure 2-1 The cognitive network of Student A

The Figure 2-2 shows the structures of connections for Student B. Student B made a number of cognitive connections between Skills, Knowledge, Values, Identities and Epistemology codes as well, but the Student’s B network features more connections as well as connections to additional elements like the epistemology of design. This means that this student tends to confirm opinions of others during the assignment solution. Student B shows strong connection between the skill of data and the epistemology of design, the skill of collaboration and the epistemology of design, the skill of design and the epistemology of design, the skill of collaboration and the skill of design and the identity of supportive and the epistemology of design.

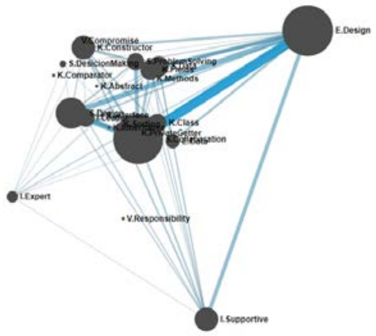


Figure 2-2 The cognitive network of Student B

Because of the fixed node positions ENA can also construct a subtracted network to enable the two networks comparison. To do this ENA subtracts the weight of each connection from one network to another and then visualizes the difference. Darker and thicker lines indicate larger differences in connections and the lighter and thinner lines indicate smaller differences in connections. The red lines indicate stronger connection for the Student A and blue lines indicate stronger connection for the Student B. Figure 2-3 shows the differences between the two networks.

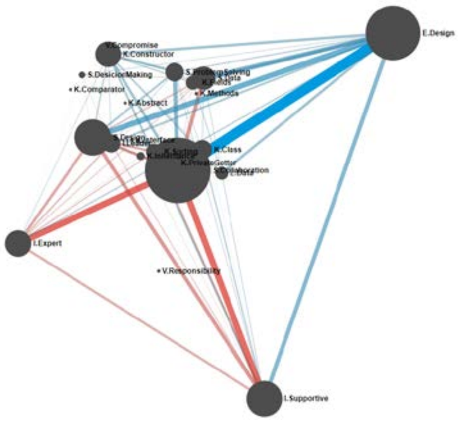


Figure 2-3 The comparison network for Student A and Student B

Student A has stronger connections between the identity of expert and the skill of collaboration, the identity of supportive and the skill of collaboration, the identity of supportive and the skill of design, the skill of collaboration and the skill of data and the identity of expert and the skill of design. Student B has stronger connections between the skill of data and the epistemology of design, the skill of collaboration and the epistemology of design, the skill of design and the epistemology of design, the skill of collaboration and the skill of design and the identity of supportive and the epistemology of design. In order to see the numeric difference between them we put the cursor above each edge in ENA Webkit Platform. For example the numeric difference between Student A and Student B for the connection between the identity of supportive and the skill of collaboration is 0.485/1 stronger for Student A. There is also another way to compare to networks which is the position of their centroids which are the means for each categorization. The furthest away they are the more different they are (Figure 2-4).

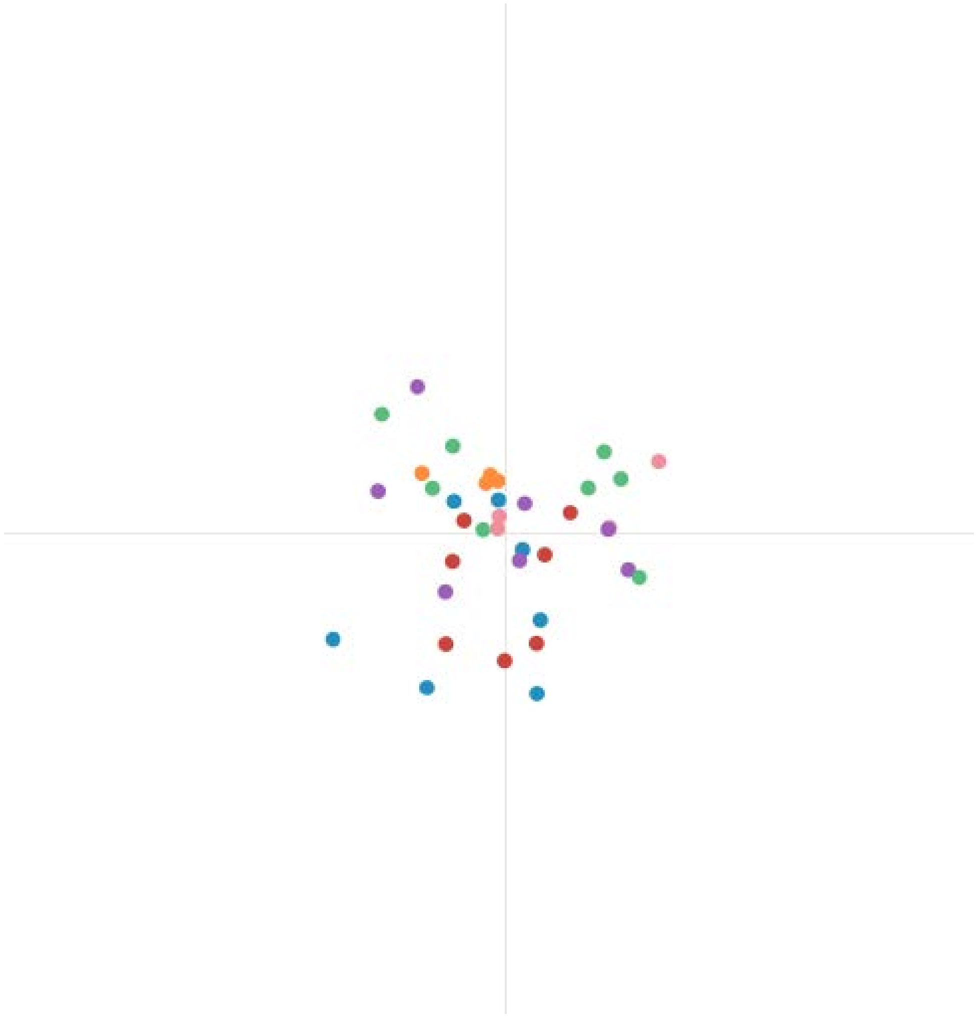


Figure 2-4 The comparison of the students based on the positions of the centroids

Literature review – Theoretical background

2.5 Introduction

ENA approach is often uses for educational or learning analysis based basically on the collaborative learning. As concluded productive learning interactions do not occur when collaborative groups have no support or scaffolding. Another main component of learning is the Interdependence of the student observed but also the jigsaw which is a puzzle consisting of a picture, cut into various pieces of different shapes that has to be fitted together. These three components are techniques that have been researched and proven effective for promoting collaborative learning. According to Johnson et al. (Johnson et al., 1993) “positive interdependence is successfully structured when group members perceive that they are linked with each other in a way that one cannot succeed unless everyone succeeds”. Thus interdependence encourages collaborations and equal contribution from the members so everyone works towards the group’s success.

The main data analyzed by ENA models comes from either collaborative games or written messages between the members. Each message or discourse can be either question, response, denial, agreement or an argument. Other basic characteristic of the data researched is social interaction and communication. E.N.A. is used mainly to analyze network models online learning, collaborative learning or problem solving. Shaffer D. (Shaffer D. et al, 2009) was the first to use E.N.A. to analyze in depth video games and the future assessment alongside with progressive education using computers. In a research paper of his he researched digital simulations for improving education using artificial teaching environments.

Shaffer D.W. was one of the first to write about the revolution of data and data analysis concerning education and social studies (Shaffer, D. W. et al, 2017). He along with his team integrates big data, data-mining, learning science, utterances analysis, cognition, statistics and ethnography into a human science. The information according to him was the combination of data and their meaning which was discovered with the data analysis. The data analysis can calculate and compare the given data and its statistics and it is able to measure the strength of their pattern using computers to handle even Big Data. He analyses the term Quantitative Ethnography as a research method to understand the meaning behind the different types of data and how to make sense out of them.

In the following sections we summarize and present the literature review for ENA organizing the studies in four different categories. We categorized the papers based on their subject of study. The first category investigates the application of ENA utilized to explore students’ connections made in different fields of Education. The second one investigates the application of ENA in Teachers’ education and training programs. The third one investigates the Application of ENA in the nutrition field; and last one investigates the application of ENA in the Medical field. We summarized the results in different tables, the columns of which are the title, the authors, the year of publication, the participants, the duration of each research and the number of sessions of each one.

Regarding the literature review from the total publications that were analyzed, 28 related to the study of applications of ENA in students and pupils, 7 related to the study of applications of ENA in teacher education, 1 related to other applications of ENA and 4 related to the application of ENA in the medical field. In the field of applications in education for students the 2 of the publications referred to middle/ high school students, 13 to university students, 8 to college and university students, 2 to students in general without referring to education level, 1 to visualization tools for educations and 2 to gaze coordination. in the field of applications in education for teachers 2 of the publications referred to teacher education in china, 3 to student-teacher internship and 2 to teachers categorization (advance - novice). In the medical field 2 referred to doctors categorization (advance - novice) and 2 to medical team observation. As for the other applications of ENA the 1 publication referred to nutrition elements and the connections between them.

2.6 Learning Analysis for Education

In this subsection we review the application of ENA in Learning Analysis for education. The participants of the researches of these papers were students from primary school to university. Table 3-1 summarizes the findings concerning the use of ENA in the analysis of educational data. In many studies, ENA was used to analyze the process of virtual internships. In most cases, the participants were from the Engineering departments of Universities and they used the tools Nephrotex (https://www.virtualinterns.org/virtual-internships/nephrotex/) and RescuShell (https://www.virtualinterns.org/virtual-internships/rescushell/) for virtual internships. In RescuShell, students designed the bottom half of an exoskeleton suit for rescue workers. This tool gives access to resources such as internal technical documents, background research, and realistic research reports. Through all this, the students learned about the inputs involved in designing exoskeletons. In Nephrotex, students were tasked with designing a nanotechnology-based membrane used for kidney dialysis. This tool gives access to internal technical documents, background research, and research reports based on real experimental data. Using this, the students developed and tested hypotheses in the provided design space.

Table 3-1: Applications of ENA in Learning Analysis for Education

Title	**Author(s)**	**Year of publication**	**Participants**	**Duration**	**Number of Sessions for Each Participant**
Epistemic network analysis: A prototype for 21st-century assessment of learning.	Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., Frank, K., Rupp, A.A., Mislevy, R.	2009	-	80 Hours	-
Modeling learning progressions in epistemic games with epistemic network analysis: Principles for data analysis and generation.	Rupp, A. A., Choi, Y., Gushta, M., Mislevy, R., Bagley, E., Nash, P., Hatfield, D., Svarovsky, G., Shaffer, D. W.	2009	3 Groups (the number of participants is not mentioned)	-	87 Time Slices
Exploring complex engineering learning over time with epistemic network analysis.	Svarovsky, G. N.	2011	1585 Schools through 2 Major Projects Project Lead the Way and the Infinity Project(they offer engineering	4 Weeks	4 Sessions one for Eck Week

			curriculum packages for middle and high school students)		
Nephrotex: Measuring First-year Students' Ways of Professional Thinking in a Virtual Internship	Arastoopour, G., &Chesler, N. C., & D'Angelo, C. M., & Shaffer, D. W., &Opgenorth, J. W., &Reardan, C. B., Haggerty, N. P., &Lepak, C. G.	2012	120 Students	-	1 Session
Epistemic trajectories: Mentoring in a game design practicum.	Nash, P., & Shaffer, D. W.	2013	5 Mentors, 3 Teams of 7 students	2 Months	4 Sessions
Measuring social identity development in epistemic games.	Arastoopour, G., & Shaffer, D. W.	2013	268 Participants	-	Data from Three Different Studies
Design of a professional practice simulator for educating and motivating first-year engineering students.	Chesler, N. C., ArastoopourIrgens, G., D'angelo, C. M., Bagley, E. A., & Shaffer, D. W.	2013	120 Students	13 Week (26 Contact-hour) Semester	1 Session, 1 Pre-Interview and 1 Post-Interview
Epistemic networks for epistemic commitments.	Knight, S., Arastoopour, G., Williamson Shaffer, DW.., Buckingham Shum, S., & Littleton, K.	2014	6 Students formed two teams of three	1 Hour for Each Team	1 Session
A novel	Chesler, N. C.,	2015	50 Students	-	2 Sessions

paradigm for engineering education: Virtual internships with individualized mentoring and assessment of engineering thinking. Journal of biomechanical engineering,	Ruis, A. R., Collier, W., Swiecki, Z., Arastoopour, G., & Williamson Shaffer, D.W.				
Look together: Analyzing gaze coordination with epistemic network analysis.	Andrist, S., Collier, W., Gleicher, M., Mutlu, B., & Shaffer, D.W.	2015	26 Participants in Dyads	-	26 Sessions 1 for each participant
Teaching and Assessing Engineering Design Thinking with Virtual Internships and Epistemic Network Analysis.	Arastoopour, G., Shaffer, D. W., Swiecki, Z., Ruis, A. R., &Chesler, N. C.	2015	46 Students	7 Weeks	2 Virtual Internships
The right kind of telling: An analysis of feedback and learning in a journalism epistemic game.	Hatfield, D.	2015	13 Undergraduate Students	16 Weeks	10 Assignments
Epistemic network analysis as a tool for engineering design assessment.	Arastoopour, G., Chesler, N. C., Shaffer, D. W., &Swiecki, Z.	2015	46 Participants	8-Week Internship	-

Gaze Mechanisms for Situated Interaction with Embodied Agents.	Andrist, S.	2016	-	-	-
Local versus global connection making in discourse.	Collier, W., Ruis, A., & Shaffer, D. W.	2016	265 Students	-	-
Teaching and assessing engineering design thinking with virtual internships and epistemic network analysis	Arastoopour, G., Shaffer, D. W., Swiecki, Z., Ruis, A. R., &Chesler, N. C.	2016	46 Students	7 Weeks	1 Session
Network analysis of interactions between students and an instructor during design meetings.	Fisher, K. Q., Hirshfield, L., Siebert-Evenstone, A., Irgens, G. A., &Koretsky, M.	2016	27 Teams of 3 Students and 2 Coaches	Each Session Lasted 30 minutes	27 Sessions for the 2 Coaches
Gaining Insight by Transforming between Temporal Representations of Human Interaction	Lund, K., Quignard, M., & Shaffer, D. W.	2017	9 Researchers	-	4 Sessions
In search of conversational grain size: modelling semantic structure using moving stanza windows.	Siebert-Evenstone, A. L., Irgens, G. A., Collier, W., Swiecki, Z., Ruis, A. R., & Shaffer, D. W.	2017(a)	44 Students	10 Week Internship	1 Session with 17 Different Activities

In search of conversational grain size: modelling semantic structure using moving stanza windows.	Siebert-Evenstone, A. L., Irgens, G. A., Collier, W., Swiecki, Z., Ruis, A. R., & Shaffer, D. W.	2017(b)	44 Students (only 5 Students Analyzed)	15 Hours	1 Session focused only on 11/17 Different Activities
Epistemic network analysis: A worked example of theory-based learning analytics.	Shaffer, D.W, &Ruis, A.	2017	311 Students	-	1 Session
Collaborative and individual scientific reasoning of pre-service teachers: New insights through epistemic network analysis (ENA)	Csanadi, A., Eagan, B., Shaffer, D. W., Kollar, I., & Fischer, F.	2017	76 Participants	-	1 Session
A network analytic approach to gaze coordination during a collaborative task	Andrist, S., Ruis, A. R., & Shaffer, D. W.	2018	26 Participants	-	26 Sessions 1 for each participant
A method for determining the extent of recent temporal context in analyses of complex, collaborative thinking.	Ruis, A., Siebert-Evenstone, A., Pozen, R., Eagan, B. R., & Shaffer, D. W.	2018(a)	652 Participants	15 Hours	1 Session

Toward a taxonomy of team performance visualization tools.	Swiecki, Z., & Shaffer, D. W.	2018	41 Papers and 48 Performance Visualization Tools	-	-
Using epistemic network analysis to examine discourse and scientific practice during a collaborative game	Bressler, D. M., Bodzin, A. M., Eagan, B., &Tabatabai, S.	2019	35 Groups of students	-	Three Game Chapters
A network-based analytic approach to uncovering the relationship between social and cognitive presences in communities of inquiry	Rolim, V., Ferreira, R., Lins, R. D., &Gašević, D.	2019	81 Students	4 weeks	Four Different Sessions
Cause and Because: Using Epistemic Network Analysis to0 Model Causality in the Next Generation Science Standards.	Siebert-Evenstone, A., & Shaffer, D. W.	2019	A set of 208 K-12 Science standards	-	-

Hatfield, D. investigated the development of the students' skills and their way of thinking using ENA and a journalism epistemic game, science.net (Hatfield, D., 2015). The participants were 13 undergraduate students who produced stories for 10 assignments. The assignments were either in group or individual ones. The data analyzed were the learning gains for players from playing science.net and were determined by comparing participant’s responses from pre-, post-, and follow-up interviews using Verbal Analysis. The codes used for the analysis were witnessing, pressing the source, investigating, narrative storytelling, removing reader barriers, detailed description, journalistic writing, reporting concepts, reader concerns journalist as reporter and as writer, informing the public, engaging readers in the story, transparency, accuracy, personal narrative and rich details. The results showed that the participants were able to develop aspects of the journalism epistemic frame as a result of playing science.net and were still present three months later.

The conversation method measures the connection of the students based on the activity yet the moving stanza window based on the activity and the contribution of each member of the team. During another research the 50 participants were divided into two teams of 25 (Chesler, N. C., 2015). The first team participated firstly in Nephrotex (Chesler, N. C. et al, 2013) and then in RescuShell and the second team used the simulations in the opposite order. The results showed that the experience of the second virtual internship gave the students the opportunity to acquire more advance engineering thinking and increased the student’s satisfaction. Also the students that took into account customer requirements were highly valued by their teammates.

The quality of the designs during the virtual internships was also analyzed using ENA (Arastoopour, G. et al, 2015a) (Arastoopour, G. et al, 2015b). The tool used in this case was Nephrotex and there were 46 participants that formed teams of 3-5 students. In this case the codes used were problem definition, planning, management, information gathering, feasibility analysis and evaluation, selection-decision and documentation. The results showed that the students that made high-quality devices talked more about managing their decision making and planning. Also it was observed that the higher Integrated Management score indicate more connections between management and other design attributes but it does not differ between the more connections between management and other design attributes.

There are cases though that ENA was used for not only monitoring the theoretical learning experience but also the practical one. An example of this use was researched by Arastoopour et al. to investigate virtual internships of engineer students to teach engineering design thinking (Arastoopour, G. et al, 2016). The internship lasted 7 weeks and the participants were 46 students and they were divided into 10 teams. In order to do this they used ENA to measure the complex thinking based on discourse analysis and the final designs produced. To carry out the study the participants were divided into groups and each group was categorized into high scoring and low scoring based on the quality of the devices that they made. The results revealed that the discourse of the teams of high-quality devices featured more connections with codes that concerned management talk and cognitive elements of engineering design than the other category. This internship gave to the student the opportunity to engage in engineering design and adopt the thinking of professional engineers. In subsequent research they studied the connections between the codes for the two categories. Students of the low performance teams had stronger connections between off-task topics, while students of high performance category made more connections between on-task topics.

Fisher, K. Q. et al used ENA to investigate the feedback between an instructor and small groups of students by analyzing the cognitive networks created by ENA. The feedback is considered one of the most important factors for successful learning (Fisher, K. Q. et al, 2016). The feedback was given by two coaches who led 27 students in total, who formed groups of three, for an undergraduate chemical, biological and environmental engineering program. The role given to the student was the one of process development engineer. Each group could pick between two reactor projects the Virtual Chemical Vapor Deposition and the Virtual Bioreactor. Each coaching session lasted 30 minutes and during them the students provided feedback about the project and the coaches guided the students to improve their strategy. The codes used for the analysis were the Student Engineering Objectives (Input Parameters, Measurement Strategy, Performance Metrics), the Coaching Objectives (Experimental Design and Strategy, Kinetics, Transport, Professional Skills, Project Contextualization) and the Stages (Surveying, Probing, Guiding, Confirmation). The comparison of the feedback was done based on the two different coaches. The results showed that there was statistically significant difference between the two coaches. The Coach 1 allowed students to choose the topic of the meeting yet coach 2 focused coaching sessions around core technical content and fundamental material.

Another one of the factors analyzed in the field of education and collaborative learning is the emotions of the students. Lund K. et al (Lund K. et al, 2017) used audio data to analyze an educational discussion for its low-level emotional indicators based on the emotions appeared in it. They analyzed the corpus JauneFluo (Fluorescent Yellow) to observe the emotional character of verbal sequences. In order to represent the temporal relationships between the emotions observed they used the tool FRIEZE, which provides researchers with a graphical representation of annotations in a temporal context. The annotation types observed were the prosody and body talk, remarkable syntactic constructions ,interactivity, non-verbal vocal productions, discourse markers, lexicon first and second person markers, macrosyntactic segments, repeated segments, turn construction units ,discourse and commitment. They analyze the data in four phases. Firstly they visualize the distribution of the indicators of emotion along with the phases. The second part is the appliance of ENA on the data to present the connections between the emotions and the different phases. The third part is a transcription of the original data to explain their aggregation. The last part was the investigation of the relationships between the observed elements by the use of a representation.

The utterances of students during the learning process can be analyzed not only based on the content of them but also based on the ways in which students make connections across turns of talk. Siebert-Evenstone, A. L. et al (Siebert-Evenstone, A. L.et al, 2017a) (Siebert-Evenstone, A. L.et al, 2017b) have analyzed the optimal ways to segment data for analysis using ENA. The data for the analysis came from 44 students who participated in RescuShell which is a 10-week long engineering virtual internship. The students formed teams of five and during the internship students develop robotic legs for a mechanical exoskeleton for use by rescue personnel. There were a total of 17 different activities as a simulation of the design process. From the 17 activities the 11 of them were used as well as the codes used for the analysis were design reasoning, performance parameters, technical constraints, client and consultant requests, collaboration and data. The results only for one of the teams were analyzed. It suggested that the conversation method and the moving stanza window method found different patterns of connection-making in student discourse.

Csanadi A. et al investigated how to combine a theoretical and a methological framework and apply them to model scientific reasoning processes and optimal units for their analysis (Csanadi A. et al, 2017). According to the theoretical framework the epistemic activities that can be analyzed are problem identification, questioning for statements, hypothesis generation, evidence generation, evidence evaluation, communications and conclusions. The methological framework used in the research is ENA. The research questions set were if collaborative and individual reasoners exhibit different epistemic networks during a professional problem solution and if the networks differ from other networks based on the same data set that has been randomly resorted. There were 76 participants in total who either solved the problem individually (16 participants) or in dyads (60 participants). The codes used in the study were based on the theoretical framework mentioned above. The results showed more complex networks for the dyads than the individuals cause evidence evaluation was connected with hypothesis generation, communicating and scrutinizing, generating solutions and non-epistemic propositions. When the data randomly resorted, identical patterns regarding complexity were similar for the dyads and the individuals with the only exception that the dyads made more connections among the highly frequent codes.

There were also researchers (Ruis, A.et al, 2018a) (Arastoopour, G. et al,2012) who investigated how to minimize the need for human annotation to determine the optimal window length for stanza. They analyzed the connectivity between the units using different moving window sizes to explore its effects on the resulting models. The software used for the engineering students was the engineering simulation Nephrotex. Through this the participants interact in 4-5 member teams with their engineering advisor using an online message program. The total process lasted 15 hours and there were 652 participants (first and second year student) and 54896 lines of chat conversations were produced. The codes used for the analysis were the window size started from 1 to 13. The results showed that the optimal window size is 7 because the relevant connections were captured for more than 95% of the sample utterances.

A necessary tool for the education is the visualization tools used by teachers (Swiecki, Z., 2018). The tools were used to create a visual representation of the space of affordance relationships of a sample of papers that described team performance, visualization tools and their affordances. There were a total of 41 papers and 48 performance visualization tools. The units for the analysis were the different tools and there were four codes for the user categories (team, educator, team and educator, researcher) and for the tool characteristics (activity, knowledge, structure, semantics, dependence, temporality, member identification and details). The results showed that the tools for the educators or teams focus more on the activity of team members and underlying data and less on the team performance data. Cognitive awareness tools focus on team knowledge, underlying data, and the contributions of all the team members and social awareness tools focus on social activity and its changes.

There were also attempts to explore the modeling the creation of intersubjectivity through gaze coordination (Andrist, S. et al, 2018) (Andrist, S., 2016) (Andrist, S. et al, 2015). The participants formed 13 pairs of two and each pair was assigned a sandwich making task. One member of the team made verbal references to visible ingredients that they chose to add in the sandwich while the other one put those ingredients into a sandwich. They also provided feedback about the results. They had to do the task twice because each member had to play both roles. There were 23 ingredients in total and they had to pick 15 for each sandwich. During each task, both participants wore mobile eye-tracking glasses. The results showed that there were differences in gaze coordination patterns that lead to breakdowns and repairs. Also they observed the properties and patterns of how gaze coordination unfolds throughout an interaction sequence.

As mentioned before the main concept analyzed using ENA is learning analysis based on collaborative problem solving. This skill is of great importance for the scientific and academic world however the majority of the students seem to lack this skill due to minimum to no exposure to it. One of these kinds of games is School Scene Investigators which is a scaffolded collaborative mobile game (Bressler, D. M. et al, 2019). Each participant of the team has an independent role in the team. The conversational discourse was analyzed to determine the connections between communication responses, language style and scientific practice. The data that were used come from audio transcripts of three teams and they were coded based on five types of scientific practice, three types of responses and one language style. The questions researched were the development scientific practice during the game play and which of the elements support this development. The results showed that there was development of the research scientific practices for the participants, who at first they chose to frame their investigation and afterwards investigated the given data. The elements proven to support this were the communal language, the accepting and the discussing responses and the interpreting data.

Recently researchers have also explored the insights which were provided about the students' development of social and critical thinking skills using CoI(Community of Inquiry). This is a framework that studies how asynchronous online communication affects learning and cognitive development of students. Rolim V. et al (Rolim, V. et al, 2019) used CoI to investigate the association of the phases of cognitive presence and the indicators of social presence and the effect of instructional scaffolds in promoting cognitive presence. The data came from 6 different time periods from a master level course in the field of software engineering of 81 students in total. A total of 1747 messages were exchanged about the 14 different topics analyzed. The research had four parts the first was a presentation of a relevant paper, the second was a literature review of a relevant paper, the third one was six questions and the fourth part was the final project. The results showed that the social "climate" during the research was improved experts made more links with the cognitive codes and tried to help the other students to gain practical experience with new projects.

Science Education was another topic researched and analyzed using ENA. Siebert-Evenstone, A. L. et al (Siebert-Evenstone, A. L. et al, 2019) investigated how the connections of the models, featuring science practices and models, can identify new insights and compare the different science disciplines. There were three dimensions the Science and Engineering Practices, Crosscutting Concepts and the Disciplinary Core Ideas. As data were used 208 K-12 science standards which were categorized into three categories which are Earth and Space Sciences, Life Sciences and Physical Sciences. As unit was considered set each line of data of written content of the Next Generation Science Standards. Furthermore each performance expectation was further segmented by the specific science and engineering practice and crosscutting concept. The results showed that there was statistical difference between Physical and Life Sciences concerning the cause and effect. More specifically Physical Sciences focus on the generation of relationships and Life Sciences focuses on their explanation.

2.6.1 Educational Collaborative Games

In this subsection we discuss the studies concerning the use of Educational Collaborative Games.

A theory based approach was provided by Shaffer D.W. for the learning analysis of an educational game (Shaffer, D. W.et al, 2009, 2014, 2015, 2017)(Rupp, A. A.et al, 2009)(Nash, P. et al, 2013) (Arastoopour, G. et al, 2013). The educational game used for the research was the Land Science a game helps students learn content and practices in urban ecology and planning and develop skills, interests, and motivation. The data were chat conversation of 311 students who used Land Science who provided 44964 lines of data. They were divided into seven groups of college students (n = 155), eight groups of high school students (n = 110), and three groups of talented high school students (n=46). The codes used for the analysis are the cognitive elements of professional urban planning practice and were 24 in total. The results showed that the gifted high school students had more and stronger connections to elements of advanced urban planning thinking than the high school students but both of these categories made fewer connections than the college students between basic professional skills and advanced urban planning thinking. Using the same data there were also compared the students based on the categorization of novices and relevant experts (Collier, W.et al, 2016). The results in that case showed that there was statistically significant difference between the two groups. Specifically the novices connect other types of justifications primarily with knowledge elements yet experts connect justifications with knowledge, skills and actions, and other justification codes.

One of the collaborative games used to research collaborative learning is Digital Zoo, an engineering epistemic game that incorporated and recreated activities and practices for middle school girls was developed and implemented (Svarovsky, G. N., 2011). Pre-, post-, and follow up interviews were conducted with each player and also the conversations of the players were recorded. This resulted to a total of three hundred and fifty audio files, thirty video files, five hundred digital notebook pages, and numerous drawings, photos, and other artifacts. The codes used for the analysis were based on the theoretical frame of ENA and were Knowledge, Skills, Identity, Values and Epistemology. The results showed that the engineering identity was mostly emphasized during the initial stages, but knowledge and skills did not differ during the stages. As for the epistemology and the values they differ during the stages but without having either an upward or a downward direction.

The Collaborative Educational Games have five main characteristics (Bressler, D. M., 2019). The participants use dialogic interactions in order to obtain all the necessary resources to solve the problem (Knight, S. et al, 2014). In order to do this they have to investigate all the given data. Also the knowledge from this type of games is analyzed beyond content mastery. Moreover the usage of argumentation during the solution of the problems gives the student the chance to negotiate with the right arguments. Finally the members of the team often have independent roles which mean that they must have social interaction and discourse to socially construct knowledge and solve the problem. Also there is the concept of Jigsaw pedagogy which means that each student in a group becomes an expert on one aspect of the activity and teaches it to the other group members.

2.7 Teachers’ Learning Analysis

In this subsection we review the application of ENA in Learning Analysis in teacher education. The participants of the researches of these papers were either novices or experienced researchers, from primary school to university. Table 3-2 summarizes the findings concerning studies using ENA in the analysis of teachers’ education data.

Table 3-2: Applications of ENA in Teachers' Learning Analysis

Title	**Author(s)**	**Year of publication**	**Participants**	**Duration**	**Number of Sessions for Each Participant**
Exploring connectedness: Applying ENA to teacher knowledge.	Orrill, C. H., & Shaffer, D. W.	2012	3 Teachers	-	1Session
Exploring coherence in teacher knowledge using epistemic network analysis	Orrill, C., Shaffer, D. W., & Burke, J.	2013	7 Teachers	90 Minutes	1 Session
Supporting teachers' intervention in students' virtual collaboration using a network based model.	Herder, T., Swiecki, Z., Fougt, S. S., Tamborg, A. L., Allsopp, B. B., Shaffer, D. W., &Misfeldt, M.	2018	3 Teachers	-	1 Session
Epistemic network analysis of students' longer written assignments as formative/summative evaluation.	Fougt, S. S., Siebert-Evenstone, A., Eagan, B., Tabatabai, S., &Misfeldt, M.	2018	16 Student Teachers	-	-

Student Teachers' Discourse During Puppetry-Based Microteaching	Wakimoto, T., Sasaki, H., Hirayama, R., Mochizuki, T., Eagan, B., Yuki, N., Funaoi, H., Kubota, Y., Suzuki, H., Kato, H.	2019(a)	36 Student Teachers	7 Hours	2 Sessions
Exploring primary school teachers' technological pedagogical content knowledge (TPACK) in online collaborative discourse: An epistemic network analysis	Zhang, S., Liu, Q., &Cai, Z.	2019	All primary and secondary school teachers in China	120 Hours	Three Sessions
Effects of Perspective-Taking Through Tangible Puppetry in Microteaching and Reflection on the Role-Play with 3D Animation.	Wakimoto, T., Sasaki, H., Hirayama, R., Mochizuki, T., Eagan, B., Yuki, N., Funaoi, H., Kubota, Y., Suzuki, H., Kato, H.	2019(b)	30 Student Teachers	15 Hours	Three Sessions

Orrill, C. et al explored the teachers' knowledge and abilities using data acquainted by interviews (Orrill, C. et al, 2013, 2012). To explore the teachers' knowledge they collected data from 3 teachers that varied mathematical ability based on their interview. The data analyzed were from their responses on the interview. The codes used for the analysis were from five categories ratio concepts, fraction concepts, representation, problem solving and other mathematical ideas. The results showed that there was significant difference between the three teachers. Specifically the expert of the three relied on ratio reasoning and problem solving skills to solve the problems, the second stronger relied to ratio understandings, but introduced fewer key ideas together and the weakest of them relied on fraction rules and operations to make sense of these concept.

Herder T. et al also investigate the concept of virtual internships for teaching internships (Herder T. et al, 2018). This was a Design-Based Research project about developing a tool that supports teachers' interventions with students in virtual internships. The software used was the Process Tab that uses a networked approach and provides information about the utterances of groups based on contributions in chat and assignments. The data analyzed came from interviews with three teachers who used the tool to provide insights about the teachers' hopes, use, and difficulties with the tool. For the analysis the students were categorized into high and low performance and their prior knowledge was used as a unit in the analysis. The results showed that the teachers found difficult to use the tool and they did not had the ambition to use because of the lack of time and energy. However the general results showed that the teachers care about viewing the created networks to evaluate performance.

In another research Fougt et al. (Fougt et al. et al., 2018) reported on exploratory trial in order to develop visualizations for 16 students' of Danish Teacher Education written assignments. The students taught a total of 66 lectures and in order to pass the lesson they had to write a five-page essay using a biographical, literary criticism, reader response, or phenomenological approach and after that relate their analysis to it. For the analysis 16 papers were used from the students who completed the essay. For this analysis two sets of keywords along with ENA were used. The main questions that were investigated are the prediction of the quality of longer student assignments and the main features for each performer category, of low, middle, and high. The results revealed that students that had more connections with description generated medium and high papers and also the high performing students had more and different connections between the codes. The type of connections also seen to be important cause middle and high performing students feature networks with more balanced connections.

ENA was also used along with specific framework to analyze teacher knowledge in technology integration practices. One of these frameworks is TPACK. There are three basic elements categories technology, pedagogy and content. The framework was used in 2013 from the Ministry of Education of China which launched a five-year teacher training program (Zhang, S., Liu et al, 2019). The participants were all primary and secondary school teachers in China who had to participate in this educational program for 120 hours. A total of 934 primary teachers participated and they categorized in 12 online learning communities based on the subjects. The data used were 561 comments from which 395 were replies. The training had three parts. Firstly they watched video cases and they solved similar ones. After that they had to participate in online discourse. For the first and the second part the results were calculated automatically by the training platform. The third part was teaching and the evaluation was given by two subject experts with more than 20 years of Chinese teaching experience based on the content, the methods and the technology application. The results showed that the pedagogical content knowledge had stronger connections with general pedagogical knowledge and weaker with the technological content knowledge. Also it was proven that the younger the participants, the greater the knowledge of technology and less in the field of education, while for the older ones the opposite proved to be the case.

In the field of education Wakimoto et al. used ENA to investigate how puppetry-based tabletop microteaching systems can contribute to student teacher training compared with normal microteaching (Wakimoto, T. et al, 2019a). To implement this they analyzed student-teachers' discourse using a puppetry based microteaching system called "EduceBoard". There were 36 participants that were undergraduate students for elementary school teacher's license. The students had to prepare a teaching plan and relative materials. The students formed 12 groups of three and each one of them had 10 minutes of microteaching. Through the analysis, the main factors were found which were improvisational thinking, situational thinking, multidimensional thinking, contextualized thinking, and reflective thinking for frameworks. The results not only showed that it improved the practical knowledge of the students' teachers but also gave them the opportunity to develop their students' learning and lead the classroom with no problems.

Puppetry was also used to examine the perspective-taking to learn a variety of pupils' viewpoints based on their reactions in undergraduate teacher education (Wakimoto, T. et al, 2019b). The participants were 30 undergraduate student teachers studying for the elementary school teacher's license and they formed groups of three. The microteaching session had two parts a role-play and a mutual feedback discussion for reflection. The codes used for the analysis were teacher management, teacher instruction, student management and student instruction. As units were considered all lines of data associated with a session ID, a group ID and student ID. The results showed that the participants tended to discuss how they should teach pupils who showed unexpected reactions in the puppetry microteaching in the 3D animation yet the participants in the video-only condition discussed how to use utterances of pupils and how they should ask pupils to do something.

2.8 Other Applications of ENA

In this subsection we review the application of ENA in Learning Analysis in the field of nutrition (Table 3-3). The analysis in this case investigates written nutrition definitions and the connections between them.

Table 3-3: Applications of ENA in the field of Nutrition

Title	Author(s)	Year of publication	Participants	Duration	Number of Sessions for Each Participant
Toward a Historical Definition of Nutrition	Ruis, A. R.	2016	-	-	-

Ruis, A. R. et al used ENA to investigate the connections between the definitions of nutrition elements (Ruis, A. R. et al, 2016). This concept had two major challenges the volume of the data and the collection and the usage of data. For the analysis they used also conceptual networks to the analyze relationship between the concepts and the language that denote them. The units chosen for the analysis where the unique sources and the codes were from three different categories: the physiological elements, adaptive elements and ecological elements. The results were based in the chronological period mentioned in each paragraph. For the period of 1800-1869 there was a larger physiological concept, for the period of 1870-1929 there were more physiological and adaptive elements and finally for the period 1930-1999 there were more holistic, balancing, physiological, adaptive and ecological elements.

2.9 Application of ENA for Medical Analysis

In this subsection we review the application of ENA in the Medical field. The participants of the researches of these papers were either novices or experienced doctors. Table 3-4 summarizes the findings of using ENA in the medical data analysis.

Table 3-4: Applications of ENA in Medical Analysis

Title	Author(s)	Year of publication	Participants	Duration	Number of Sessions for Each Participant
The hands and head of a surgeon: Modeling operative competency with multimodal epistemic network analysis.	Ruis, A. R., Rosser, A. A., Quandt-Walle, C., Nathwani, J. N., Shaffer, D. W., & Pugh, C. M.	2018	40 Participants	15 minutes	1 Session
Using epistemic network analysis to identify targets for educational interventions in trauma team communication.	Sullivan S., Warner-Hillard C., Eagan B. , Thompson R. , Ruis A. R. , Haines K. , Pugh C.M. , Shaffer D.W., Jung H.S.	2018	80 Participants	-	-
Quantifying the qualitative with epistemic network analysis: a human factors case study of	Wooldridge, A. R., Carayon, P., Shaffer, D.	2018	4 Participants	7 Workdays for 15 hours	7 Workdays

task-allocation communication in a primary care team	W., & Eagan, B				
Multiple uses for procedural simulators in continuing medical education contexts	Ruis, A. R., Rosser, A. A., Nathwani, J. N., Beems, M. V., Jung, S. A., & Pugh, C. M	2019	58 Participants	6 Hours	2 Sessions

One concept researched by Ruis A.R. et al. (Ruis A.R. et al., 2018) was a method to access intraoperative performance by modeling how surgeons integrate psychomotor, procedural, and cognitive skills to manage errors. There were 45 participants that were general surgery residents from seven different hospitals. All the participants had to perform the final steps of a laparoscopic ventral hernia (LVH) repair using a physical, box-style simulator. The data collected were audio and video data from performing a simulated laparoscopic ventral hernia repair. The codes used for the analysis were frustration, identifying errors, operative planning, giving instructions and six types of errors. The participants analyzed had to commit at least one error during the operation. There were finally 40 participants that were separated into two groups the high performing (n=20) and the low performing group (n=20). The results showed that high performing participants had more connections to Operative Planning after committing errors and between Identifying Errors and Operative Planning. The low performing participants made more often operative plans because of Motor Errors and they more likely reacted to errors with frustration.

Apart from surgical teams, trauma teams were analyzed too using ENA. Sullivan S. investigated the ability of the members of trauma teams to use nontechnical skills like collaboration to complete tasks (Sullivan, S., 2018). For the utterances categorization Verbal Response Modes (VRM) was used to categorize based on the relationship created by what is said. The 8 categories were disclosure, edification, advisement, confirmation, question, acknowledgement, interpretation and reflection. Each utterance was coded twice once for the form and once for the intent. Sixteen teams of 5 participants participated in the training simulations. Each team consisted of a trauma chief, a surgeon, emergency medicine resident and 2 emergency medicine nurses. There were 8 different trauma scenarios that were randomly assigned to the teams. The teams also were divided to low performing and high performing. The results showed that the lower performing team featured more and stronger connections to edification, disclosure, and interpretation yet high performing teams had stronger connections to acknowledgement, advisement, and confirmation.

Human Factors and Ergonomics was another topic analyzed using ENA (Wooldridge, A. R., 2018). It is a qualitative type of analysis to provide feedback for the work systems and processes. In this case ENA was used to quantify qualitative data and analyze the communication in a primary care team. The participants were a team of four people a physician, a nurse, a medical assistant and a unit clerk. The data were collected by a researcher who observed the team for 15 hours over 7 workdays and they collected a total of 83 task-allocation communications. The results show that physician and unit clerk were most successful allocating tasks, the unit clerk is the only team member to successfully allocate tasks to the physician. Also it is observed that the nurse is the least successful assigning tasks and that the physician successfully allocates tasks to the unit clerk more often than the nurse.

One of the fields that ENA was applied to is the one of procedural simulations. Ruis A.R. et al (Ruis A.R. et al., 2019) researched whether procedural simulations can help advance learners in the medical field to develop beyond learning and rehearsing basic procedural knowledge and skills. The networks were formed from the medical participants that were divided on the basis of their experience in the medical field into advance and novice learners. There were 58 participants in total and they had to participate in a six-hour course consisted of a two-hour lecture and a four-hour practicum. For the observation of the participants all the sessions were audio and video recorded and four researchers participated in the observation process. The codes used were Mesh Repair, General Anatomy, Pathological Anatomy, Requesting Advice, Troubleshooting and Real World Case. The results showed that not only the experts but also novices were helped too. The novices used the simulators in the traditional sense to obtain basic knowledge for the procedures while advance learners used it to study adaptation to different clinical presentations.

3 Case Study

3.1 Introduction

This chapter presents the practical application of the ENA described in detail above. The field in which ENA was applied is the scientific framework of Object-Oriented Programming (OOP).Specifically we used ENA to analyze chat to see if we can detect differences between the connections made by Computer Science undergraduate students with different performance levels in an Object-Oriented Programming (OOP) course and their scores in a collaborative solving OOP assignment. We also investigated how the knowledge and skills on core cognitive elements of OOP can affect their performance. For the purpose of the research we conducted an experiment where the participants were asked to solve a collaborative assignment of OOP. From this process we extracted and collected the dialogues to analyze them using ENA.

The following sections explain the procedure of the research and all the different scenarios that were analyzed using ENA. Also the steps followed are presented, as well as the results of the analysis.

3.2 Methodology

To participate in the research students had to fill in a form where they declared their grades in 3 courses: Procedural Programming, Data Structures and Object-Oriented Programming. The invitation was mainly for students who should have been taught Object Oriented Programming and thus 10 groups were selected where 37 students participated. The participants formed groups of three or four members by themselves and they were asked to solve an exercise in Java using the Eclipse Programming Platform for coding and the Zoom Meeting Software for Communication. Zoom was used because of the feature of remote control because the members of each group were apart due to Covid-19. They only communicated using written text messages in Zoom.

The participants in the research were rewarded according to their participation both in the solution of the assignment and their collaboration with 1.5 points. There were 7 groups of four students and three of three students.

Students were told that they would be participating in an assignment that involved a collaborative exercise solving and discussion in groups. They were also instructed that their assignment description would be downloaded by the University Digital Platform

Open E-Class (https://openeclass.uom.gr) and that they had to log in to Specific Zoom Meeting Rooms. There were initially three different supervisors as hosts. Five minutes before the beginning of the solution each group was asked to pick a host from the members. After that the supervisors left the meeting. This decision was made because if the students were to be supervised by someone the messages would not be impulsive and the students would hesitate due to the presence of the supervisor. This decision was proven right because none of the chats had hesitation or unnecessary formalities but also the students communicated quickly and effectively.

For the solution of the assignment one of the students opened the Programming Development Platform Eclipse and started to write code based on personal knowledge and skills and with the help of the rest of the members. Other members of the group could ask for permission to write code using Remote control. In this way they could write on the same platform without being on the same PC. They communicated with written messages only with Zoom but also sent written code from similar exercises that they had found online or solved before.

The coding assignment concerned fundamental concepts of OOP: Classes, fields, methods, constructors; Subclasses, inheritance, method overriding, relationship between subtyping and inheritance; Abstract classes; Visibility of class members; Comparators; Sorting methods. The assignment had the following scenario: "A chain of stores that sells toys (Greek-made or imported) and children's books of Greek publishing houses decided to implement a simple application for the management of its product data." In the first phase the assignment required the calculation of the final price of the products sold by the chain and the printing tags. They had to declare and develop the parent class Product and the subclasses Book and Game. To calculate the final price they had to know if the game was imported or not and for the books they had to know the discount percentage. Then they had to display the sorted list of all games and books according to a) name and b) to their price (in ascending order) and display the following statistics: average book price, average game price, and most expensive game price. They had 90 minutes to solve the exercise and 15 minutes to upload the exercise and the meeting's chat on the LMS Open E-Class.

In order to investigate Research Questions concerning different types of students' and groups' profiles based on their pre and post assignment performance or the use of fundamental concepts of OOP, we considered the following two cases:

Case 1:

- Based on their Performance in both relevant course and the solution of the assessment.

Students were divided into 4 categories based on their performance in the course of object-oriented programming and their performance in the assignment they solved. The total lines of discourse collected were 2800. The students were divided based on their grade point average in the course which was 7.2 / 10 and based on the exercise grade point average which was 7.8 / 10. Based on these, the following 4 categories were created:

- High-to-High: Students with grade higher than 7.2 and assignment score higher than 7.8
- High-to-Low: Students with grade higher than 7.2 and assignment score less than 7.8
- Low-to-High: Students with grade less than 7.2 and assignment score higher than 7.8
- Low-to-Low: Students with grade less than 7.2 and assignment score less than 7.8

Of the total number of students, 40.54% belong to the High-to-High category (15/37), 16.21% belong to the High-to-Low category (6/37), 27.02% belong to the Low-to-High category (7/37) and 16.21% belong to the Low-to-Low category (6/37). Students with a score above or below the mean course grade and groups with scores above or below the mean assignment grade were categorized as High or Low scoring, respectively. We had 3 different categories of groups based on members' individual and group scores. There were 3 groups of High student course score and High group assignment score (High-to-High); 3 groups of High or Low student course score and High group assignment score (High/Low-to-High); and 3 groups of High or Low student course score and Low group assignment score (High/Low-to-Low).Of the total number of students, 30.30% (10/33) belong to the High-to-High category, (11/33) 33.33% belong to the High/Low-to-High category, and 36.36% (12/33) belong to the High/Low-to-Low category.

Table 4-1 presents members' individual and group scores as well as the category

Table 4-1 Members and Groups Scores

	Member1: score in	**Member2: score in OOP**	**Member3: score in OOP**	**Member4: score in OOP**	**Group score**	**Category of the Group**

	OOP course/category	course/category	course/category	course/category	in the assignment	
Group1	10/High	10/High	10/High	-	9/High	High-to-High
Group2	4/Low	4/Low	6/Low	10/High	5.5/ Low	High/Low-to-Low
Group3	4/Low	9/High	10/High	10/High	7.5/ Low	High/Low-to-Low
Group4	4/Low	4/Low	10/High	10/High	8.5/ High	High/Low-to-High
Group5	8/High	10/High	10/High	10/High	8.5/ High	High-to-High
Group6	4/Low	7/Low	8/High	10/High	5/Low	High/Low-to-Low
Group7	4/Low	5/Low	7/Low	10/High	8.5/ High	High/Low-to-High
Group8	9/High	9/High	10/High	-	8.5/ High	High-to-High
Group9	4/Low	10/High	10/High	-	10/ High	High/Low-to-High
Group10	4/Low	4/Low	6/Low	6/Low	8/High	Low-to-High

Case 2:

- Based on the use of the three fundamental Concepts of OOP as a group.

For this categorization we examined which of the groups used in their solution the following three fundamentals concepts of OOP: Abstract, Inheritance and Comparator. From the 10 groups; 2 used all the fundamental concepts of OOP (Group5 and Group9), 5 used the fundamentals Abstract and Inheritance (Group1, Group3, Group4, Group7 and Group8), 1 used only the concept of Inheritance (Group10) and 2 did not use any of the fundamental concepts of OOP (Group2 and Group6). Thereby resulting to the four following categories:

- None: Groups that did not use any of the three fundamental concepts of OOP
- Inheritance: Groups that used only the fundamental concept of Inheritance
- Abstract-Inheritance: Groups that used the fundamental concepts of Abstract and Inheritance
- Abstract-Inheritance-Comparator: Groups that used the fundamental concepts of Abstract, Inheritance and Comparator

3.2.1 Log Data

The file used for the analysis with the ENA Webkit software has as lines the utterances of the participants. Figure 4-1 presents the log data of the analysis. The columns of the file apart from the codes used for ENA are the following:

- GroupID: The group of the student whose utterances are in the specific line.
- Exersice: The number of the assignment observed. There were initially two assignments but because of the difficulty level of the second assignment and the lack of time only the first one was analyzed.
- Object-Oriented Programming: The Students Grade in the course of Object-Oriented Programming.
- Team Members: The numbers of the total members of the group.
- Student Categ: The category of the students based on the first categorization mentioned above.
- Category2: The category of the Group of the students based on the grades on the assignment and the course of OOP.
- Inh/Abs/Comp: The category of the students based on the second categorization mentioned above.
- Time: The time counting from the time the students started to solve the assessment.
- UserID: The ID Number of the User whose utterances are in the specific line.
- Chat: The utterance observed.
- Codes: The codes observed in the utterance (see next subsection for the coding scheme).

1	GroupID	Exersice	Catetgory2	Object-Oriented Programming	Exersice	Team Members	Student Categ	Inh/Abs/Comp	Time	User	Chat	Codes	K.Sorting	K.Class	K.Abstr
2	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:01:21	USER1:	jekinv	-	0	0	
3	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:01:23	USER2:	Καλησπέρα!	-	0	0	
4	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:01:37	USER3:	καλησπερα	-	0	0	
5	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:04:03	USER1:	ας μιλαμε από δω	-	0	0	
6	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:04:09	USER1:	φαινεται η οθονη ετι	-	0	0	
7	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:04:16	USER3:	ναι	-	0	0	
8	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:04:24	USER1:	ωραια	S.Collaboration,I.Leader	0	0	
9	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:04:24	USER3:	βλέπουμε το operec	-	0	0	
10	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:04:45	USER1:	επρεπε να το παιξεις	-	0	0	
11	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:04:45	USER2:	ωραία	-	0	0	
12	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:04:59	USER2:	Πρέπει να είμαστε κε	-	0	0	
13	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:05:06	USER1:	δεν νομιζω	-	0	0	
14	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:05:11	USER2:	Ήμουν για απορίες	-	0	0	
15	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:05:19	USER1:	το 2ο πιστευω	-	0	0	
16	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:05:53	USER2:	έγινε	-	0	0	
17	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:06:30	USER2:	Ωραία ανοιγουμε ολ	S.Collaboration,I.Leader	0	0	
18	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:08:30	USER1:	ειπε πως να γραψε το	-	0	0	
19	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:08:39	USER2:	Όχι	-	0	0	
20	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:08:44	USER2:	αν εννοείς να φαινετ	-	0	0	
21	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:09:00	USER2:	Εγώ διαβάζω ακομα	-	0	0	
22	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:09:17	USER3:	εγω την διαβασα	-	0	0	
23	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:09:33	USER3:	μπραβο φτιαξε τη m	K.Methods, K.Class, I.Leader	0	1	
24	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:10:00	USER1:	να φτιαξουμε κλαση	K.Class, I.Quandary	0	1	
25	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:10:01	USER3:	και σιγουρα μια κλα	K.Class,E.Design,I.Leader	0	1	
26	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:10:10	USER1:	και να κληρονομει η	S.ProblemSolving, K.Inheritance, I.Quandary	0	0	
27	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:10:17	USER3:	ναι συμφωνω	E.Design,S.Collaboration,	0	0	
28	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:10:37	USER2:	ωραία	E.Design	0	0	
29	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:10:50	USER1:	αφηνω την product	K.Class	0	1	
30	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:11:07	USER3:	η item θα εχει ονομα	K.Fields,I.Expert	0	0	
31	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:11:13	USER3:	για γνωρισματα	K.Fields,I.Expert	0	0	
32	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:11:20	USER1:	ωραια	E.Design	0	0	
33	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:11:23	USER2:	φτιαξε τους construc	K.Constructor	0	0	
34	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:12:03	USER1:	τι εχει το book?	K.Class	0	1	
35	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:12:13	USER1:	εκδοτη?	K.Fields	0	0	
36	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:12:14	USER2:	Ονομα	K.Fields,S.Design	0	0	
37	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:12:22	USER2:	τιμη	K.Fields,S.Design	0	0	
38	Group1	Exersice1	High-High	10	9	3	High-to-High	Abstract,Inheritance	0:12:30	USER3:	syggrafea	K.Fields,S.Design	0	0	

Figure 3-1 Log Data

3.2.2 Coding scheme

We coded each message of chat data using our OOP Epistemic Frame elements coding scheme, which identifies domain-specific frame elements. Our coding scheme includes a set of 23 codes relevant to OOP practice based on the ACM Computing Curricula 2020 (https://www.acm.org/education/curricula-recommendations) using Epistemic Frame Theory (D.W. Shaffer et al, 2004), (D.W. Shaffer, 2006) as a guide for students practicing on OOP. The codes represent Knowledge (fundamental concepts of OOP: class, inheritance, constructor etc.); Skills (Problem Solving, Collaboration, Design); Identity (Supportive, Leader, Expert, Quandary); Epistemology, Values (Compromise, Responsibility). In the following we give code names and their descriptions:

- Knowledge: the perceptions that people share in the community
    - K.Abstract: Knowledge and Usage of Abstract classes in Java
    - K.Comparator: Knowledge and Usage of Comparators in Java
    - K.Constructor: Knowledge and Usage of Constructors in Java
    - K.Class: Knowledge and Usage of Classes in Java
    - K.Fields: Knowledge and Usage of Fields in Java
    - K.Getter: Knowledge and Usage of Getters methods to access class fields in Java
    - K.Inheritance: Knowledge and Usage of Inheritance in Java
    - K.Interface: Knowledge and Usage of Interfaces in Java
    - K.Methods: Knowledge and Usage of Methods in Java
    - K.Sorting: Knowledge of how to sort objects in Java
- Identity: the way community members see themselves
    - I.Expert: The person that is knowledgeable about or skillful on Java
    - I.Leader: The person who “leads” the group
    - I.Quandary: The person who hesitates and asks questions
    - I.Supportive: The person who offers help to another
- Skills: the things people do in the community
    - S.Collaboration: Collaboration during the exercise solution
    - S.Data: The skill of data management
    - S.DesicionMaking: The skill of Decision Making
    - S.Design: The skill of code design
    - S.ProblemSolving: The skill of solving problems

- Epistemology/Confirmation: the guarantees that justify actions or beliefs, as legitimate, within the community
    - E.Data: The confirmation of correct data usage
    - E.Design: The confirmation for right design
- Values: the beliefs held by members of the community
    - V.Compromise: The value of someone the compromise on something
    - V.Responsibility: The value of someone to take responsibility for something

Chat was manually analyzed to assign each message to an appropriate code. The coding process was appropriately validated. We used two human raters and the interrater reliability analysis shows that all pairwise agreements among rater 1, rater 2 meet standards for Cohen's kappa.

3.2.3 Research Questions

The research questions investigated are the following:

- RQ1a. What types of connections between codes are made by each students' Group?
    - We research this question to investigate the main characteristics of each group.
- RQ1b. Is there a significant difference between the discourse networks of the ten different students' Groups?
    - Indicatively we compare Group 1 with the other Groups in the research to investigate the differences between them and see the different perspectives of each group.
- RQ2a. What types of connections between codes are made by groups in the High-to-High category? What types of connections between codes are made by Groups in the High/Low-to-High category? What types of connections between codes are made by Groups in the High/Low-to-Low category?
    - We investigate the main characteristics of each category to find the most important connections of each one.
- RQ2b. Is there a significant difference between the discourse networks of groups of the three categories: High-to-High, High/Low-to-High and High/Low-to-Low?
    - We compare the 3 categories with each other to find the main differences between them.
- RQ3. Is there a significant difference between the discourse networks of groups of the same Category?
    - We compare the groups of the same category to identify their differences and see if they are significantly different with each other.

- RQ4. Is there a significant difference between the discourse networks of the categories of the Groups based on the fundamental OOP concepts they used?
    - We compare the 4 categories to find the differences between them.

3.3 Results

3.3.1 RQ1a. What types of connections between codes are made by each student's Group?

The Group 1 consists of three members. The grade of the assignment for this group is 9/10 and the group belongs to the High-to-High category. All of its members belong to the Student Category of High-to-High and for the solution they used the fundamentals concepts of Abstract and Inheritance. The stronger connection of this group is between the Skill of Collaboration and the Skill of Design(S.Collaboration-S.Design: 0.439). There are also strong connections between the Identity of Expert and the Skills of Design (I.Expert-S.Design: 0.199) and Collaboration (I.Expert-S.Collaboration: 0.190). Lastly there are strong connections between the Epistemology of Design and the Skills of Collaboration (E.Design-S.Collaboration: 0.273) and Design (E.Design-S.Design: 0.292). The Figure 4-2 resents the cognitive network of the Group1.

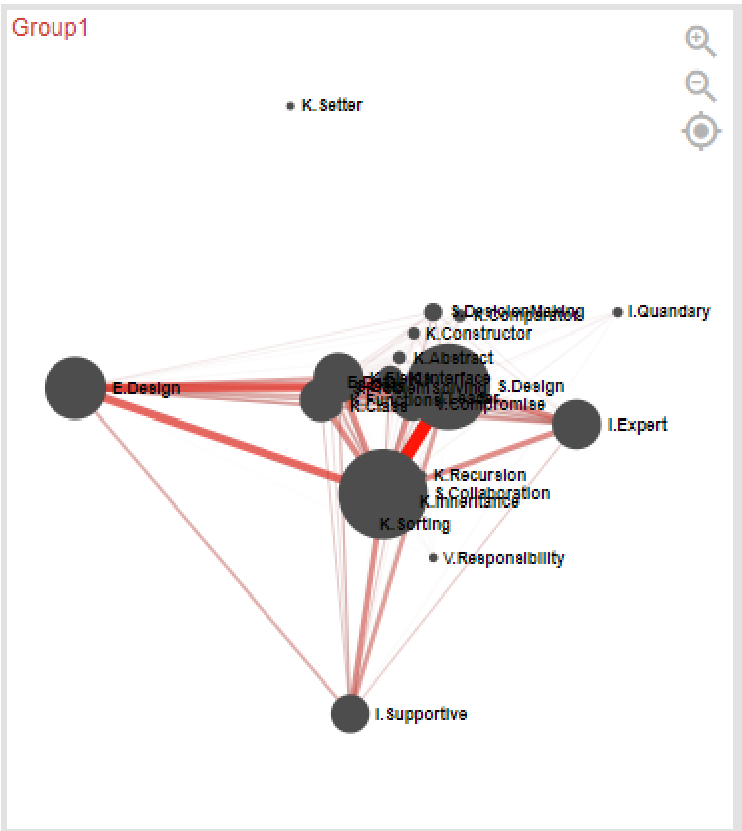


Figure 3-2 Network of the Group1

The Group 2 consists of four members. The grade of the assignment for this group is 5.5/10 and the Group belongs to the High/Low-to-Low category. Two of its members belong to the Student Category of High-to-Low and two of them to Low-to-Low and for the solution they did not use any of the three fundamentals concepts. The stronger connection of this group is between the Skill of Collaboration and the Skill of Design (S.Collaboration-S.Design: 0.393). There are also strong connections between the Identity of Expert and the Skills of Design (I.Expert-S.Design: 0.135) and Collaboration (I.Expert-S.Collaboration: 0.182). Lastly there are strong connections between the Epistemology of Design and the Skills of Collaboration (E.Design-S.Collaboration: 0.373) and Design (E.Design-S.Design: 0.281). The Figure 4-3 resents the cognitive network of the Group2.

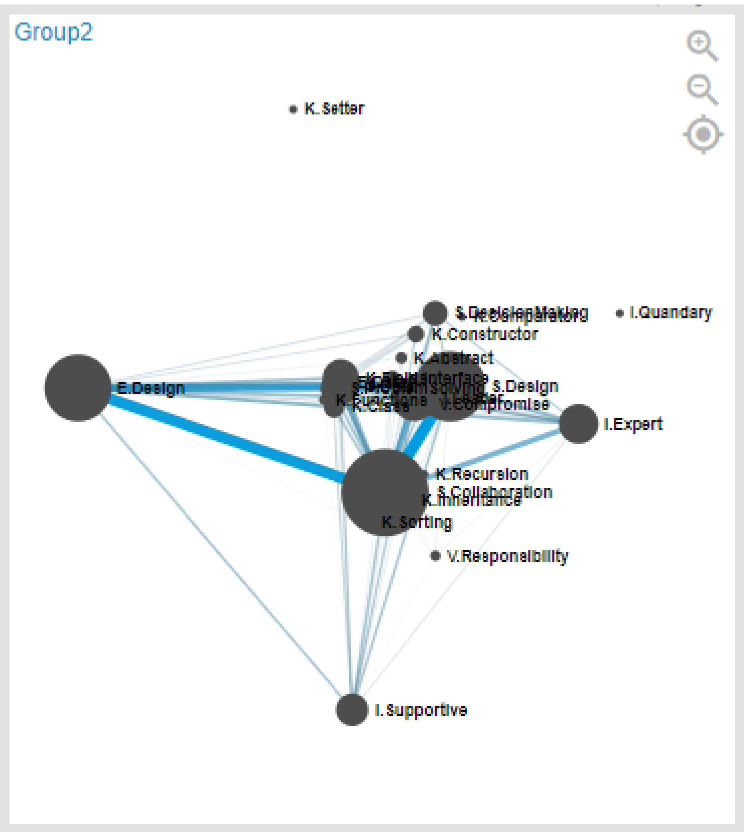


Figure 3-3 Network of the Group2

The Group 3 consists of four members. The grade of the assignment for this group is 7.5/10 and the Group belongs to the High/Low-to-Low category. Three of its members belong to the Student Category of High-to-Low and one of them to Low-to-Low and for the solution they used the fundamentals concepts of Abstract and Inheritance. The stronger connection of this group is between the Skill of Collaboration and the Skill of Design (S.Collaboration-S.Design: 0.394). There are also strong connections between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.260), between the Identity of Leader and the Skill of Collaboration (I.Leader-S.Collaboration: 0.258)and between the Identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.318). Lastly there are strong connections between the Epistemology of Design and the Skills of Collaboration (E.Design-S.Collaboration: 0.387) and Design (E.Design-S.Design: 0.216). The Figure 4-4 resents the cognitive network of the Group3.

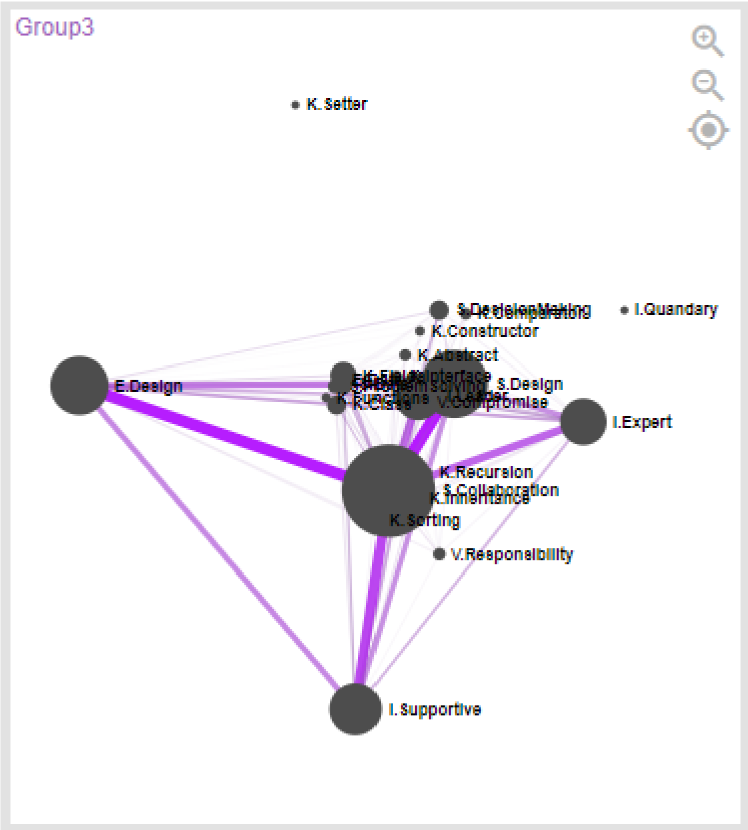


Figure 3-4 Network of the Group3

The Group 4 consists of four members. The grade of the assignment for this group is 8.5/10 and the Group belongs to the High/Low-to-High category. Two of its members belong to the Student Category of High-to-Low and two of them to Low-to-Low and for the solution they used the fundamentals concepts of Abstract and Inheritance. The stronger connection of this group is between the Skill of Collaboration and the Skill of Design (S.Collaboration-S.Design: 0.428). There are also strong connections between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.277), between the Identity of Supportive and the Skill of Design (I.Supportive-S.Design: 0.249) and between the Identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.371). Lastly there are strong connections between the Epistemology of Design and the Skills of Collaboration (E.Design-S.Collaboration: 0.345) and Design (E.Design-S.Design: 0.241). The Figure 4-5 resents the cognitive network of the Group4.

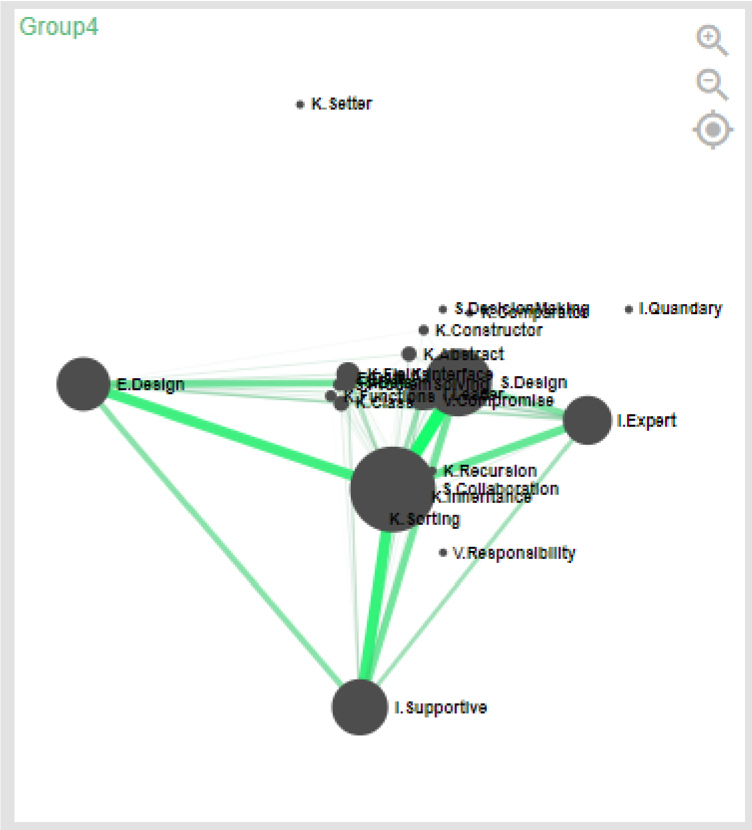


Figure 3-5 Network of the Group4

The Group 5 consists of four members. The grade of the assignment for this group is 8.5/10 and the Group belongs to the High-to-High category. All of its members belong to the Student Category of High-to-High and for the solution they used the fundamentals concepts of Abstract, Inheritance and Comparator. The stronger connection of this group is between the Skill of Collaboration and the Skill of Design (S.Collaboration-S.Design: 0.459). There are also strong connections between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.301), between the Identity of Supportive and the Skill of Design (I.Supportive -S.Design: 0.241) and between the Identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.362). Lastly there are strong connections between the Epistemology of Design and the Skills of Collaboration (E.Design-S.Collaboration: 0.231) and between the Identity of Leader and the Skill of Collaboration (I.Leader-S.Collaboration: 0.319). The Figure 4-6 resents the cognitive network of the Group5.

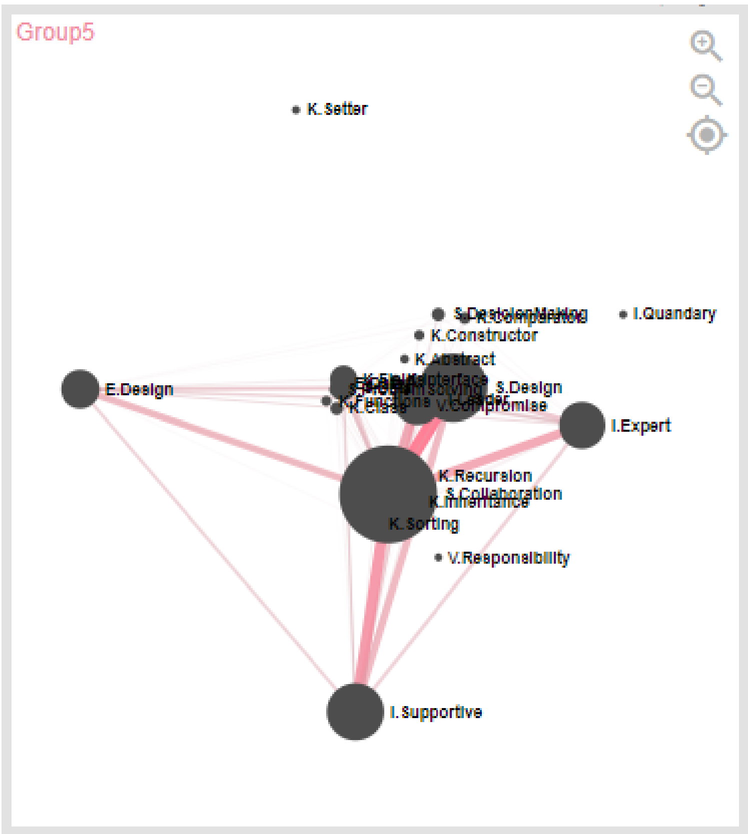


Figure 3-6 Network of the Group5

The Group 6 consists of four members. The grade of the assignment for this group is 5/10 and the Group belongs to the High/Low-to-Low category. Two of its members belong to the Student Category of High-to-Low and two of them to Low-to-Low and for the solution they did not use any of the three fundamentals concepts. The stronger connection of this group is between the Identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.471). There are also strong connections between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.228), between the Skill of Collaboration and the Skill of Design (S.Collaboration - S.Design: 0.437) and between the Epistemology of Design and the Skill of Collaboration (E.Design-S.Collaboration: 0.369). Lastly there are strong connections between the Identity of Supportive and the Skill of Design (I.Supportive-S.Design: 0.256). The Figure 4-7 resents the cognitive network of the Group6.

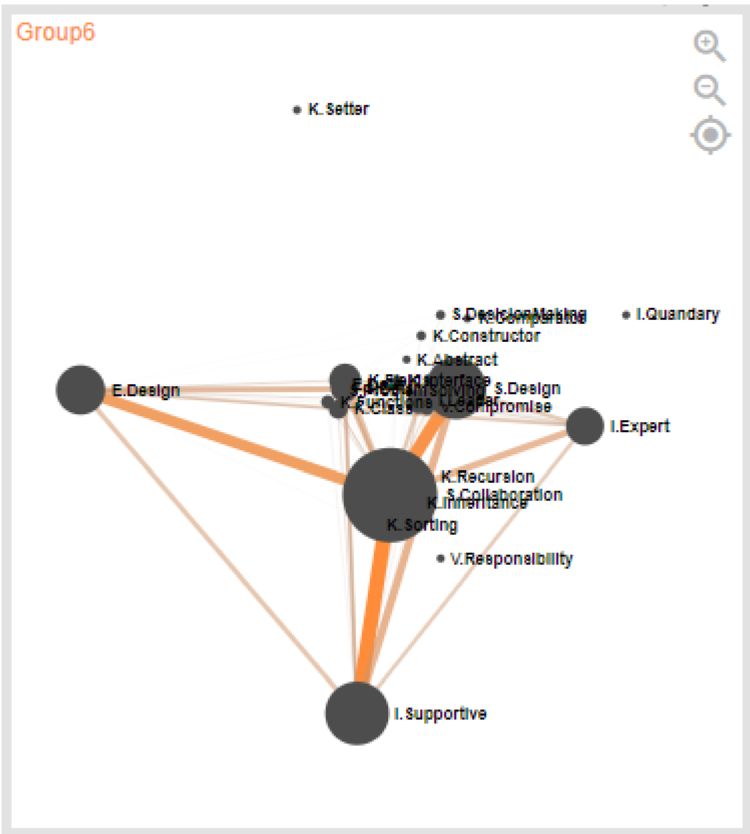


Figure 3-7 Network of the Group6

The Group 7 consists of four members. The grade of the assignment for this group is 7.5/10 and the Group belongs to the High/Low-to-High category. Three of its members belong to the Student Category of Low-to-Low and one of them to High-to-Low and for the solution they used the fundamentals concepts of Abstract and Inheritance. The stronger connection of this group is between the Skill of Collaboration and the Skill of Design (S.Collaboration-S.Design: 0.447). There are also strong connections between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.332) and between the Identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.321). Lastly there are strong connections between the Epistemology of Design and the Skills of Collaboration (E.Design-S.Collaboration: 0.361) and Design (E.Design-S.Design: 0.207). The Figure 4-8 resents the cognitive network of the Group7.

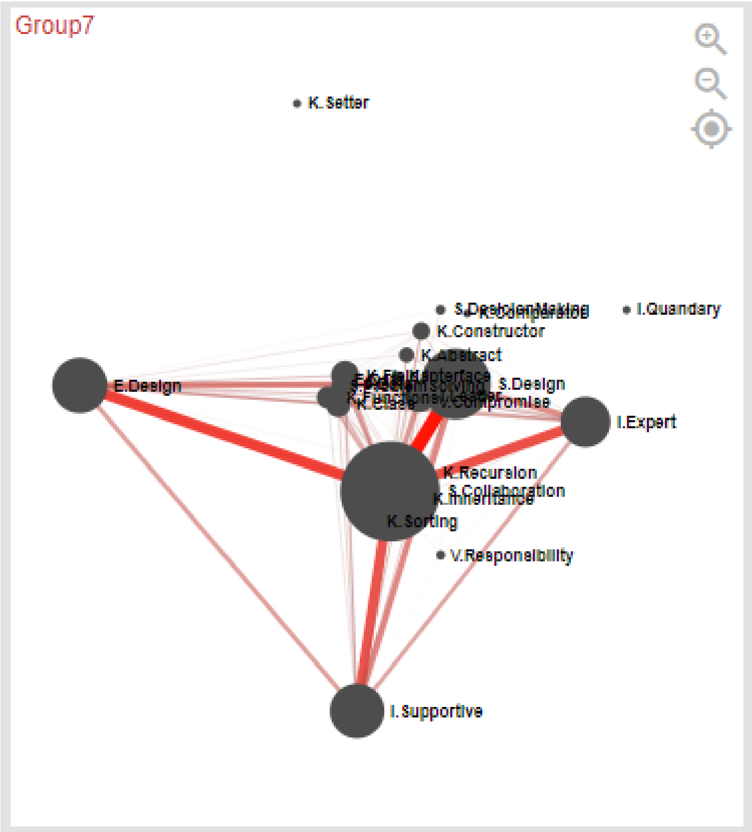


Figure 3-8 Network of the Group7

The Group 8 consists of three members. The grade of the assignment for this group is 8.5/10 and the Group belongs to the High-to-High category. All of its members belong to the Student Category of High-to-High and for the solution they used the fundamentals concepts of Abstract and Inheritance. The stronger connection of this group is between the Skill of Collaboration and the Skill of Design (S.Collaboration-S.Design: 0.440). There are also strong connections between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.214), between the Identity of Supportive and the Skill of Design (I.Supportive -S.Design: 0.254) and between the Identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.388). Lastly there are strong connections between the Epistemology of Design and the Skills of Collaboration (E.Design-S.Collaboration: 0.332) and Design (E.Design-S.Design: 0.217). The Figure 4-9 resents the cognitive network of the Group8.

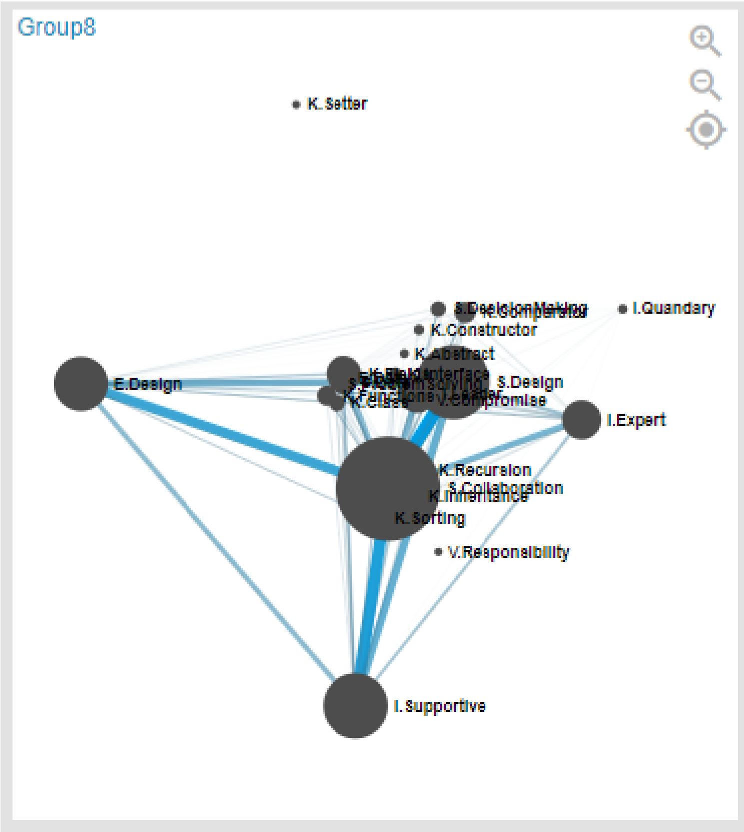


Figure 3-9 Network of the Group8

The Group 9 consists of three members. The grade of the assignment for this group is 10/10 and the Group belongs to the High/Low-to-High category. Two of its members belong to the Student Category of High-to-High and one of them to Low-to-High and for the solution they did not use any of the three fundamentals concepts. The stronger connection of this group is between the Identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.458). There are also strong connections between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.244), between the Skill of Collaboration and the Skill of Design (S.Collaboration - S.Design: 0.323) and between the Epistemology of Design and the Skill of Collaboration (E.Design-S.Collaboration: 0.277). Lastly there are strong connections between the Identity of Leader and the Skill of Collaboration (I.Leader-S.Collaboration: 0.275). The Figure 4-10 resents the cognitive network of the Group9.

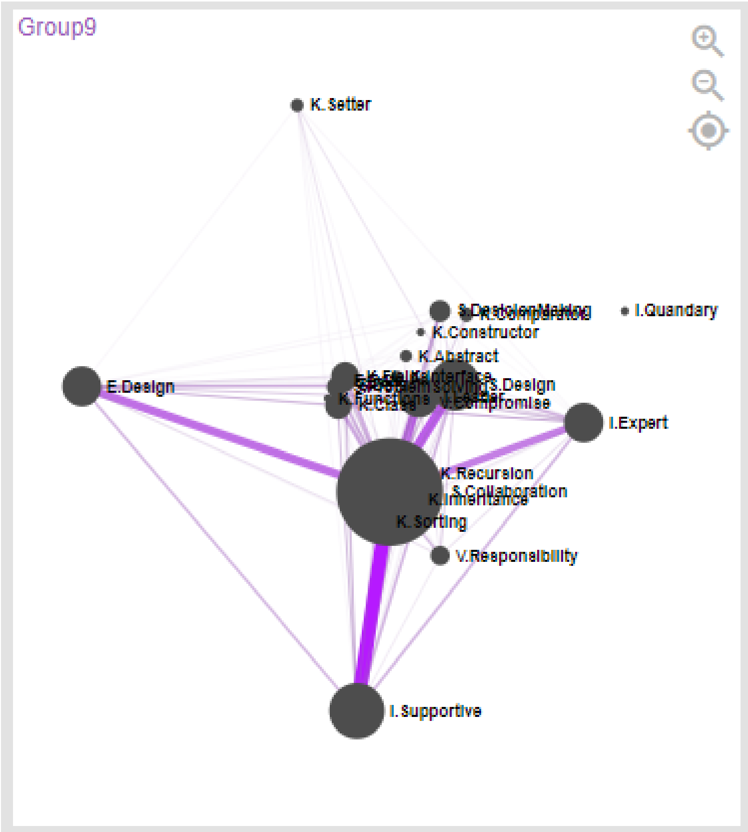


Figure 3-10 Network of the Group9

The Group 10 consists of four members. The grade of the assignment for this group is 8/10 and the Group belongs to the Low-to-High category, which it is not analyzed because it consists only from one group. All of its members belong to the Student Category of High-to-Low and for the solution they used the fundamentals concept of Inheritance. The stronger connection of this group is between the Skill of Collaboration and the Identity of Supportive (S.Collaboration-I.Supportive: 0.475). There are also strong connections between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.297), between the Identity of Supportive and the Skill of Design (I.Supportive -S.Design: 0.251) and between the Skill of Design and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.441). Lastly there is strong connection between the Epistemology of Design and the Skills of Collaboration (E.Design-S.Collaboration: 0.242). The Figure 4-11 resents the cognitive network of the Group10.

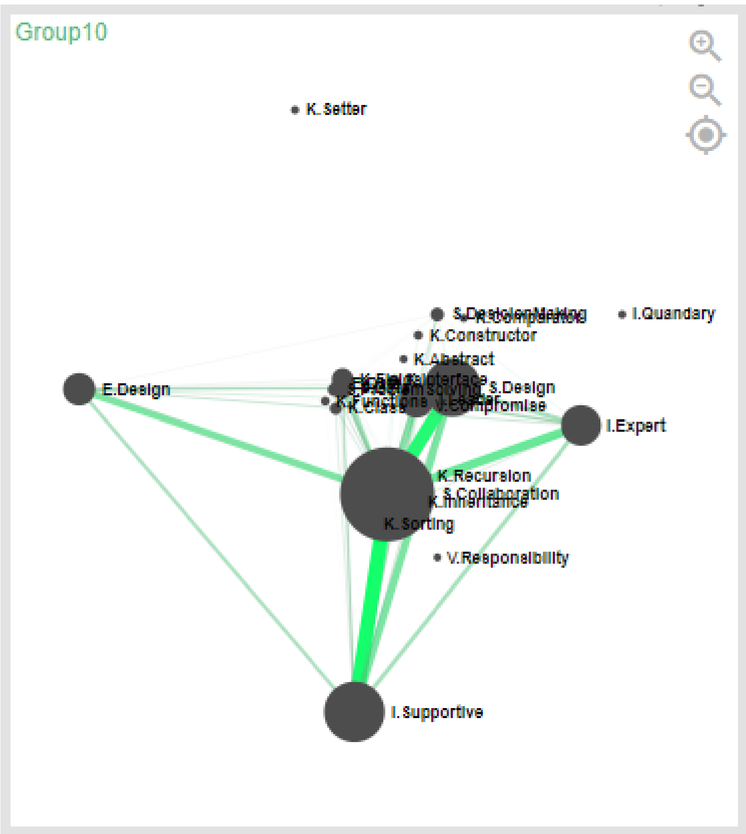


Figure 3-11 Network of the Group10

3.3.2 RQ1b. Is there a significant difference between the discourse networks of different students' Groups?

For this research question we compare the Group 1 with the other Groups of the Participants in the research.

In order to investigate RQ1b we research the ability of the students to solve an Object Oriented Exercise using Java based the groups that they created by themselves.

The units, the codes and the conversation that were chosen are the following:

- Units: GroupID, UserID
- Conversation: GroupID, UserID,Student.Categ
- Stanza Window: Moving Window Consisted of 4 Lines
- Codes: All the codes given in the file
- Comparison: GroupID

The figure we see in ENA WebKit, after selecting the parameters, is the following (Figure 4-12), where we can see the centroids and the confidence intervals for ten student groups

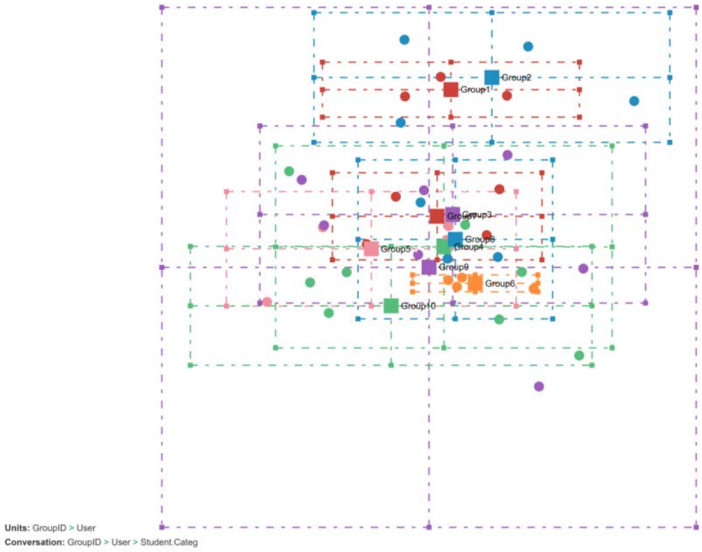


Figure 3-12 Centroids of the Ten Groups

By default, the ENA space in which the centroids are displayed is determined by the first (x) and second (y) dimensions, the dimensions that represent the largest data variation. The numbers in parentheses next to the axis labels indicate the percentage of variation in the data relating to these dimensions. In this case, the dimension x (SVD1) represents 20.3% of the variance in the data and the dimension y (SVD2) represents 18%.

As shown in Figure 4-12 the groups feature some differences. To determine the difference more accurately, we can perform an independent samples t - test. To do this, simply select the two samples we want to compare from the drop-down menu on the left, in the tab "Stats". When we do this, we will see averages for the two samples, along with the t-score, the p value and Cohen's d, a measure of the magnitude of the effect. In order to analyze them further we will present the differences of the connections between the two groups. This difference is calculated by subtracting the weight of the edge of one network from the other. The numbers in the brackets is the numeric differences of each edge between the two Groups or Categories. The research question that is answered in the following paragraphs is:

RQ1b. Is there a significant difference between the discourse networks of different students' Groups?

For this case study we are going to compare each of the other groups with the Group1 and find the important differences between them. The Group1 was chosen for the analysis because it was further of the majority of the other thus the majority of the comparisons would find the difference between this and the other Groups significantly different. The numbers in the brackets is the numeric differences of each edge between the two Groups or Categories.

3.3.2.1 Comparison Group1 and Group2

In this case, the difference on the first and the second dimension is not significantly different:

- *Along the X axis, a two sample t test assuming unequal variance showed Group1 (mean=0.196, SD=0.884, N=3 was not statistically significantly different at the alpha=0.05 level from Group2 (mean=0.896, SD=1.911, N=4; t(4.417)= -0.646, p=0.550, Cohen's d=0.442).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group1 (mean=2.228, SD=0.188, N=3 was not statistically significantly different at the alpha=0.05 level from Group2 (mean=2.435, SD=0.694, N=4; t(3.567)= -0.569, p=0.603, Cohen's d=0.376).*

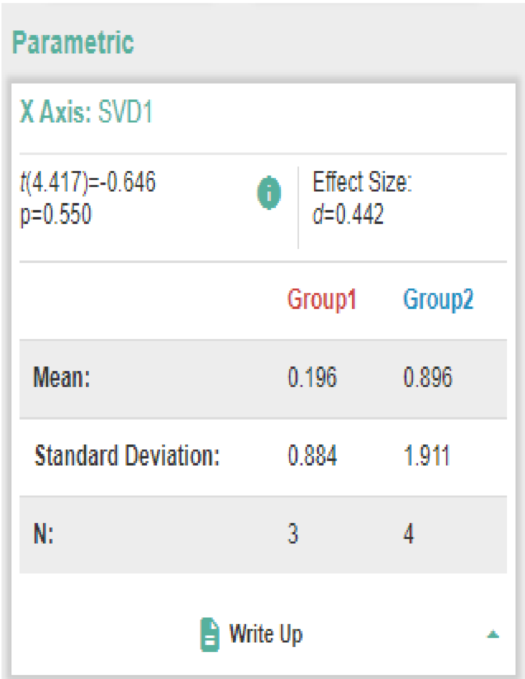


Figure 3-13 The results of the Independent T-Test for the first dimension

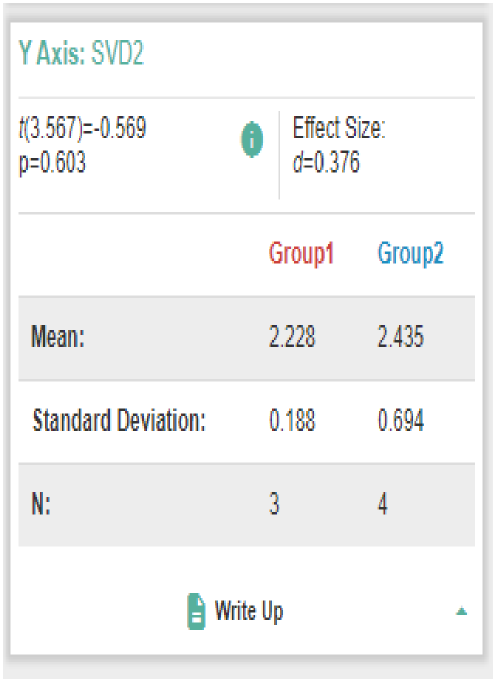


Figure 3-14 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-13 and Figure 4-14 show the independent t-test for the first and second dimension for Group1 and Group2. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are close to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved. The Figure 4-15 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R.

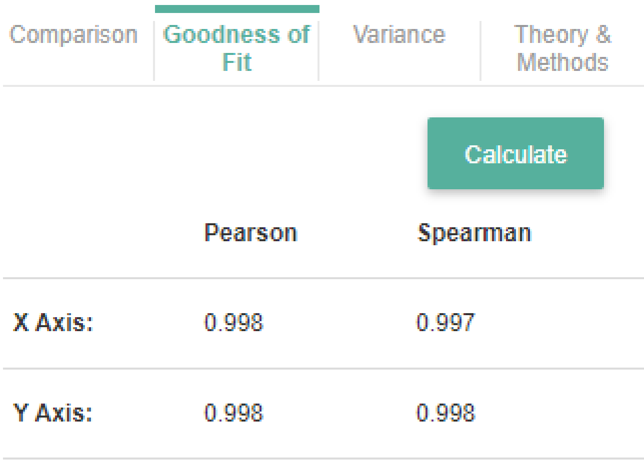


Figure 3-15 Pearson's and Spearman's R for the Goodness of Fit

To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two groups the two following networks for the two groups will appear. The Figure 4-16 shows the cognitive network for the Group1, the Figure 4-17 shows the Comparison Networks for Group1 and Group2 and the Figure 4-18 shows the cognitive network for the Group2.

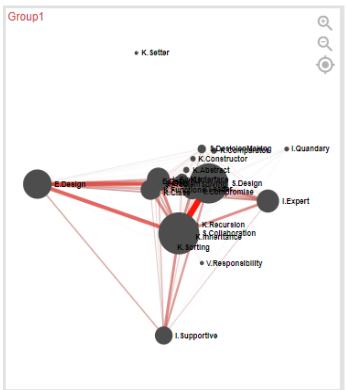


Figure 3-16 The Network of the centroid for Group1

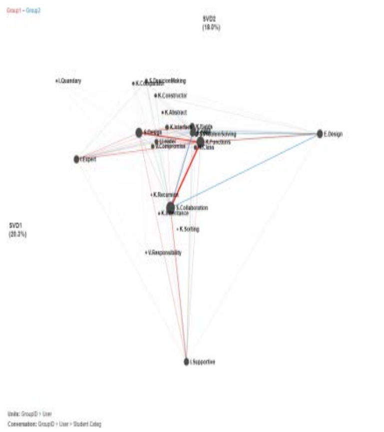


Figure 3-17 Comparison of Group1 and Group2

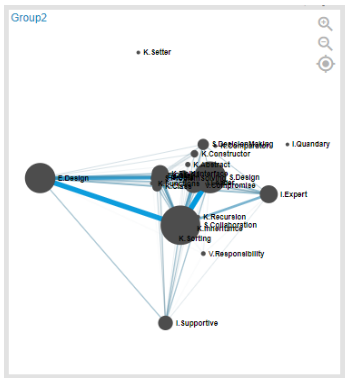


Figure 3-18 The Network of the centroid for Group2

There are little to no differences between the two groups analyzed. The differences worth presenting are:

The members of Group 1 focused more on the Methods of the OOP because there are strong connections between Knowledge of Methods and Skills of Design (K.Methods - S.Design: 0.154) and Collaboration (K.Methods - S.Collaboration: 0.192) as well as the Epistemology of Design (K.Methods - E.Design: 0.104). There are also stronger connections with the Identities of Expert (K.Methods-I.Expert: 0.09) and Leader (K.Methods-I.Leader: 0.092). That means that the members had the knowledge and the skills needed but also there was someone that guided them during the process.

The members of the Group 2 have stronger connections with the Epistemology category codes which mean that they tend to confirm more often the answers of the other members. Specifically there were strong connections between Epistemology of Design and Skill of Collaboration (S.Collaboration - E.Design: 0.100) and Skill of Data (S.Data-E.Design: 0.075), and the Epistemology of Data (E.Data - E.Design: 0.071) but there was also strong connection between skill of Collaboration and skill of Data(S.Collaboration - S.Data: 0.107).

3.3.2.2 Comparison Group1 and Group3

In this case there is not a significant difference on the first dimension but there is a significant difference on the second one:

- *Along the X axis, a two sample t test assuming unequal variance showed Group1 (mean=0.196, SD=0.884, N=3 was not statistically significantly different at the alpha=0.05 level from Group3 (mean=0.226, SD=2.071, N=4; t(4.258)= -0.026, p=0.980, Cohen's d=0.018).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group1 (mean=2.228, SD=0.188, N=3 was statistically significantly different at the alpha=0.05 level from Group3 (mean=0.104, SD=0.947, N=4; t(3.311)= 4.373, p=0.018, Cohen's d=2.859).*

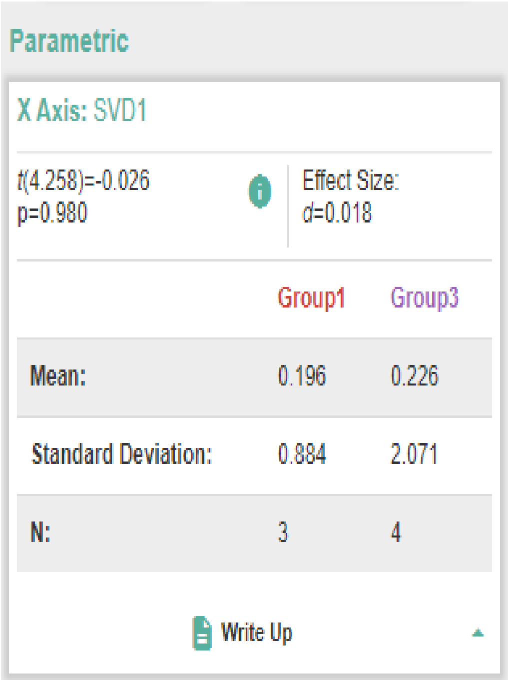


Figure 3-19 The results of the Independent T-Test for the first dimension

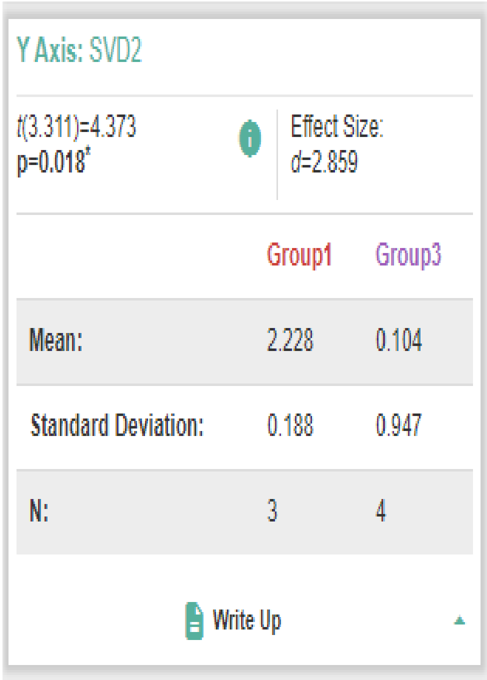


Figure 3-20 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-19 and Figure 4-20 show the independent t-test for the first and second dimension for Group1 and Group3. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are close to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved. The Figure 4-21 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R.

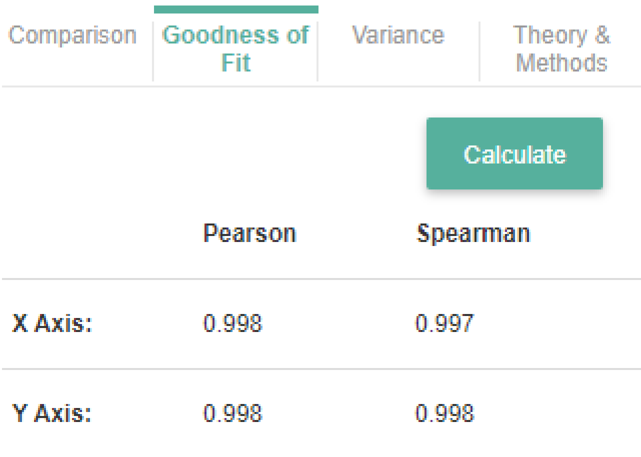


Figure 3-21 Pearson's and Spearman's R for the Goodness of Fit

To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two groups the two following networks for the two groups will appear. The Figure 4-22 shows the cognitive network for the Group1, the Figure 4-23 shows the Comparison Networks for Group1 and Group3 and the Figure 4-24 shows the cognitive network for the Group3.

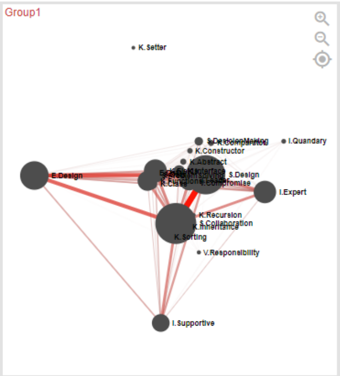


Figure 3-22 The Network of the centroid for Group1

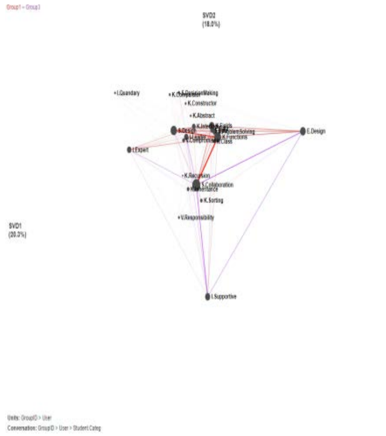


Figure 3-23 Comparison of Group1 and Group3

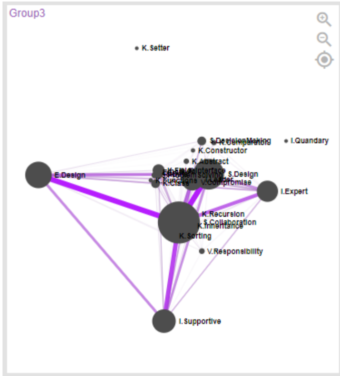


Figure 3-24 The Network of the centroid for Group3

There are many differences between the two groups analyzed. The differences worth presenting are:

The members of Group 1 focused more on Methods and Fields of data of the OOP because there are stronger connections between Knowledge of Methods and Skills of Design (K.Methods - S.Design: 0.159) and Collaboration (K.Methods - S.Collaboration: 0.191) as well as the Epistemology of Design (K.Methods - E.Design: 0.104). Also there are strong connections between Knowledge of Fields and Skills of Design (K.Fields- S.Design: 0.105) as well as the Epistemology of Design (K.Fields- E.Design: 0.088).There are also stronger connections with the identities of Expert (K.Methods-I.Expert: 0.084) and Leader (K.Methods-I.Leader: 0.096). Lastly there are strong connections between the Skill of Problem Solving, the Collaboration (S.ProblemSolving-S.Collaboration: 0.082) and the Design (S.ProblemSolving-S.Design: 0.075). That means that the members had the knowledge and the skills needed but also there was someone that guided them during the process.

The members of the Group 3 have stronger connections with the Epistemology of Design which mean that they tend to confirm more often the design ideas of others. Specifically there were strong connections between Epistemology of Design and the skills Collaboration (S.Collaboration - E.Design: 0.114) and the Identity of Supportive (I.Supportive-E.Design: 0.084). The connections from the Skill of Collaboration and the

Identities of Supportive (S.Collaboration-I.Supportive: 0.129), Leader (S.Collaboration-I.Leader: 0.078) and Expert (S.Collaboration-I.Expert: 0.07) are also strong. That means that students ofGroup1 have more advance knowledge and design skills in comparison to the third group but the members of the Group 3 have distinctive roles in the group.

3.3.2.3 Comparison Group1 and Group4

In this case there is not a significant difference on the first dimension but there is a significant difference on the second one:

- *Along the X axis, a two sample t test assuming unequal variance showed Group1 (mean=0.196, SD=0.884, N=3 was not statistically significantly different at the alpha=0.05 level from Group4 (mean=0.076, SD=1.807, N=4; t(4.529)= 0.116, p=0.913, Cohen's d=0.080).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group1 (mean=2.228, SD=0.188, N=3 was statistically significantly different at the alpha=0.05 level from Group4 (mean=-0.447, SD=1.081, N=4; t(3.240)= 4.854, p=0.014, Cohen's d=3.164).*

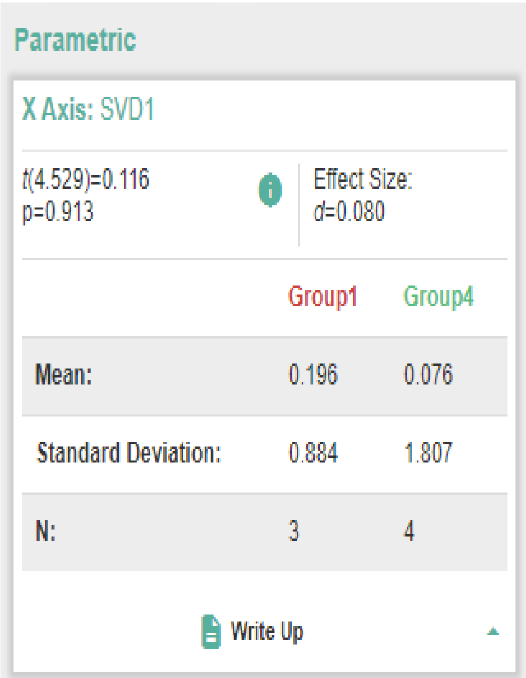


Figure 3-25 The results of the Independent T-Test for the first dimension

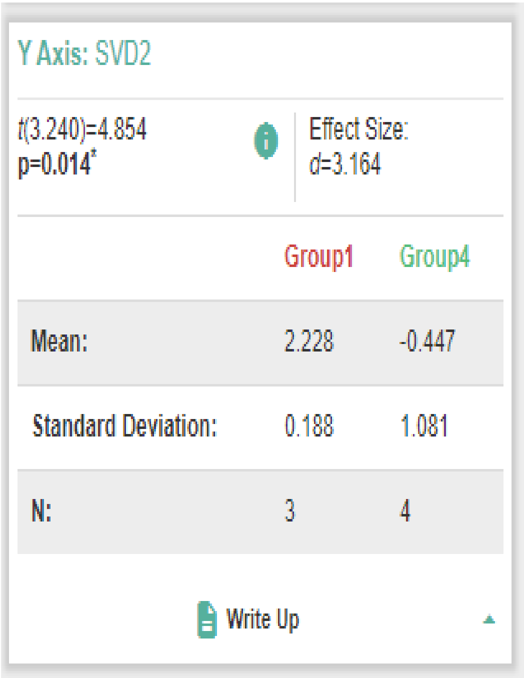


Figure 3-26 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-25 and Figure 4-26 show the independent t-test for the first and second dimension for Group1 and Group4. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are close to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved. The Figure 4-27 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R.

Comparison | **Goodness of Fit** | Variance | Theory & Methods

Calculate

	Pearson	Spearman
X Axis:	0.998	0.997
Y Axis:	0.998	0.998

Figure 3-27 Pearson's and Spearman's R for the Goodness of Fit

To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two groups the two following networks for the two groups will appear. The Figure 4-28 shows the cognitive network for the Group1, the Figure 4-29 shows the Comparison Networks for Group1 and Group4 and the Figure 4-30 shows the cognitive network for the Group4.

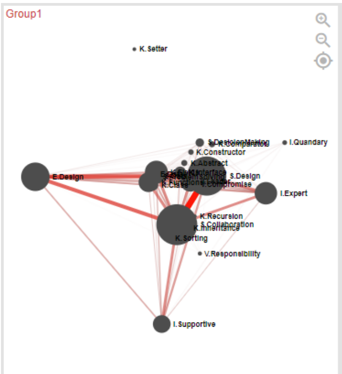


Figure 3-28 The Network of the centroid for Group1

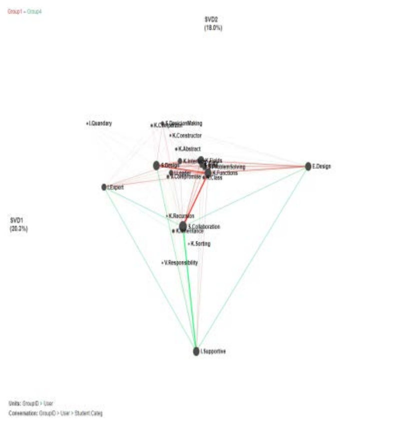


Figure 3-29 Comparison of Group1 and Group4

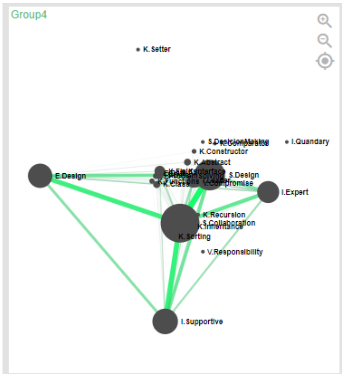


Figure 3-30 The Network of the centroid for Group4

There are many differences between the two groups analyzed. The differences worth presenting are:

The members of the Group 1 have focused more on the Methods and the Fields of the data of the OOP because there are stronger connections between Knowledge of Methods and Skills of Design (K.Methods - S.Design: 0.148) and Collaboration

(K.Methods - S.Collaboration: 0.176) as well as the Epistemology of Design (K.Methods - E.Design: 0.098). Also there are strong connections between Knowledge of Fields and the Skills of Design (K.Fields - S.Design: 0.103) as well as the Epistemology of Design (K.Fields - E.Design: 0.086).There are also stronger connections with the identities of Leader (K.Methods-I.Leader: 0.096).That means that the members had the good knowledge of the basics needed to solve the exercise and collaborated over that.

The members of the Group 4 have stronger connections with the Epistemology of Design and the identity of Supportive. Specifically there were strong connections between Epistemology of Design and the skills Collaboration (S.Collaboration - E.Design: 0.072) and the Identity of Supportive (I.Supportive -E.Design: 0.094). The connections from the identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.182), the Skill of Design (I.Supportive-S.Design: 0.097) and the identity of Expert (I.Supportive-I.Expert: 0.09). That means that students of Group 1 have more advance knowledge and skills in comparison to the other group but the members of Group 4 have distinctive roles in the group with the identity of the Supportive to appear prominent which means that the members of Group 4 helped more each other in order to solve the exercise.

3.3.2.4 Comparison Group1 and Group5

In this case there is not a significant difference on the first dimension but there is a significant difference on the second one:

- *Along the X axis, a two sample t test assuming unequal variance showed Group1 (mean=0.196, SD=0.884, N=3 was not statistically significantly different at the alpha=0.05 level from Group5 (mean=-1.164, SD=1.556, N=4; t(4.804)= 1.462, p=0.206, Cohen's d=1.024).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group1 (mean=2.228, SD=0.188, N=3 was statistically significantly different at the alpha=0.05 level from Group5 (mean=-0.481, SD=0.612, N=4; t(3.716)= 8.337, p=0.002, Cohen's d=5.539).*

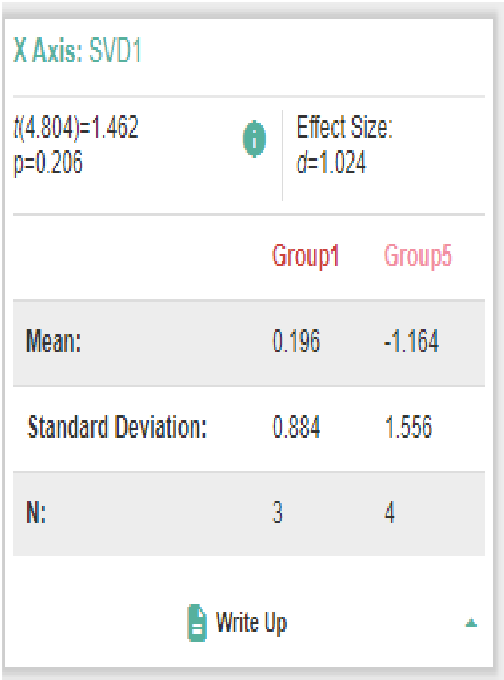


Figure 3-31 The results of the Independent T-Test for the first dimension

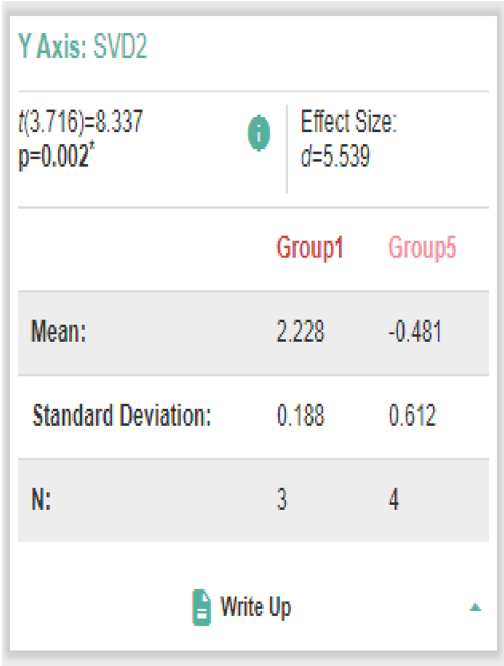


Figure 3-32 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-31 and Figure 4-32 show the independent t-test for the first and second dimension for Group1 and Group5. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are close to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved. The Figure 4-33 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R.

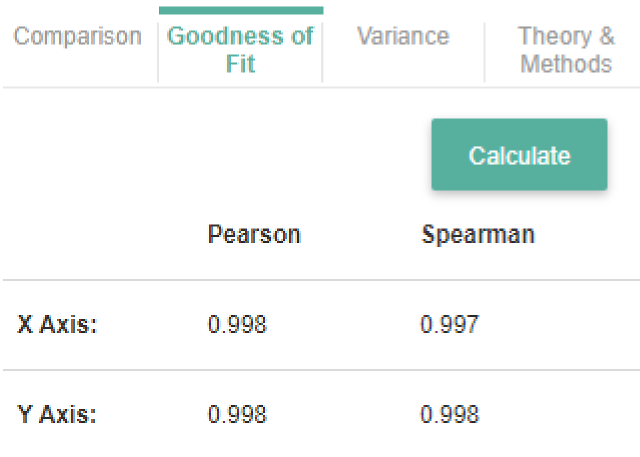


Figure 3-33 Pearson's and Spearman's R for the Goodness of Fit

To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two groups the two following networks for the two groups will appear. The Figure 4-34 shows the cognitive network for the Group1, the Figure 4-35 shows the Comparison Networks for Group1 and Group5 and the Figure 4-36 shows the cognitive network for the Group5.

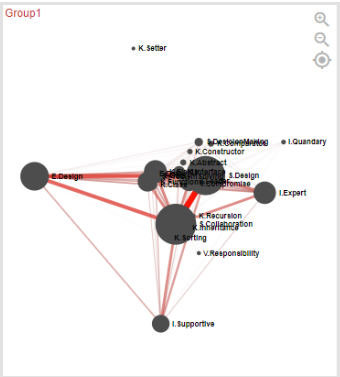


Figure 3-34 The Network of the centroid for Group1

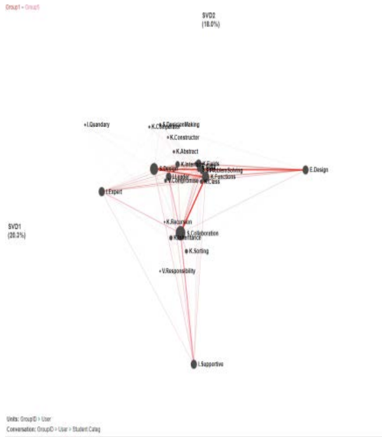


Figure 3-35 Comparison of Group1 and Group5

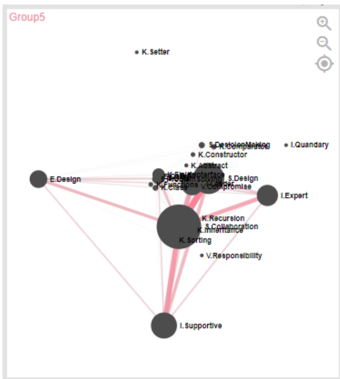


Figure 3-36 The Network of the centroid for Group5

There are many differences between the two groups analyzed. The differences worth presenting are:

The members of the Group 1 focused more on the Methods of OOP because there are strong connections between Knowledge of Methods and Skills of Design (K.Methods - S.Design: 0.155) and Collaboration (K.Methods - S.Collaboration: 0.185), and the Identity of Leader (K.Methods-I.Leader: 0.097). There are also stronger connections between the Skill of Design and the epistemology of Design(S.Design-E.Design: 0.157) as well as between the skill of Collaboration and the skill of Problem Solving(S.ProblemSolving-S.Collaboration: 0.091).That means that the members had the good knowledge of the methods needed and the Problem Solving skill to solve the exercise and collaborated over that.

The members of the Group5 have stronger connections with the Skill of Collaboration and the identity of Supportive. Specifically there were strong connections between Skill of Collaboration and the identities of Supportive (I.Supportive-S.Collaboration: 0.173), Leader (I.Leader-S.Collaboration: 0.138) and Expert (I.Expert-S.Collaboration: 0.111). The connections from the identity of Supportive and the skill of Design (I.Supportive-S.Design: 0.089) and the identity of Leader (I.Supportive-I.Leader: 0.09). That means that students of Group 1 have more advance knowledge and skills in comparison to the other group but the members of the Group 5 have distinctive roles and collaboration between the members of the group with the identity of the Supportive to appear prominent which means that the members of the Group 5 helped more each other in order to solve the exercise.

3.3.2.5 Comparison Group1 and Group6

In this case there is not a significant difference on the first dimension but there is a significant difference on the second one:

- *Along the X axis, a two sample t test assuming unequal variance showed Group1 (mean=0.196, SD=0.884, N=3 was not statistically significantly different at the alpha=0.05 level from Group6 (mean=0.611, SD=0.673, N=4; t(3.655)= -0.679, p=0.538, Cohen's d=0.543).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group1 (mean=2.228, SD=0.188, N=3 was statistically significantly different at the alpha=0.05 level from Group6 (mean=-1.067, SD=0.090, N=4; t(2.697)= 27.983, p=0.000, Cohen's d=23.850).*

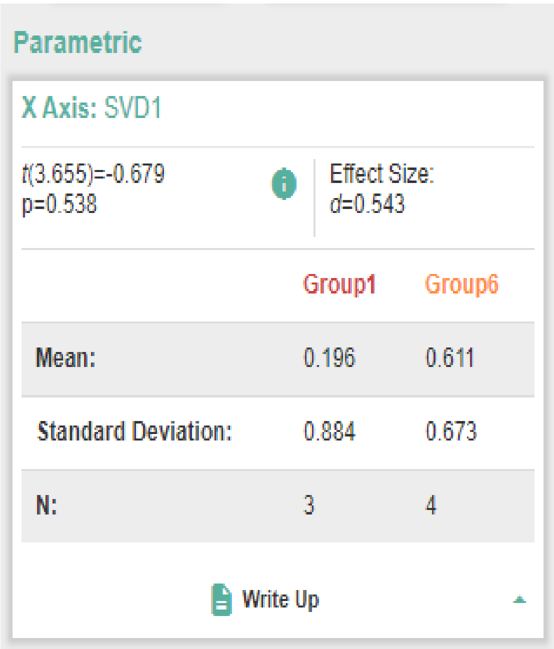


Figure 3-37 The results of the Independent T-Test for the first dimension

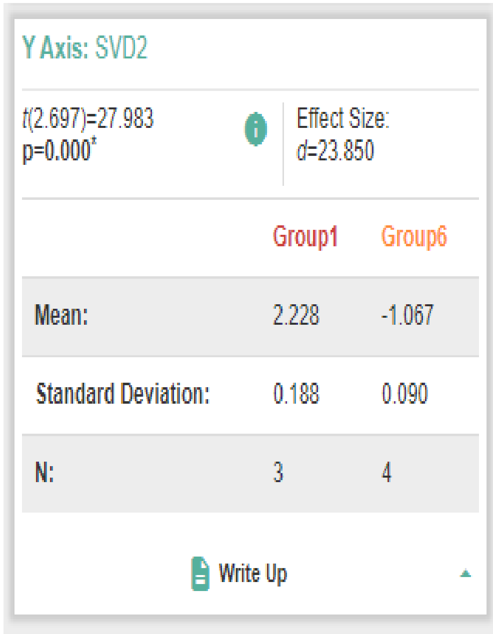


Figure 3-38 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-37 and Figure 4-38 show the independent t-test for the first and second dimension for Group1 and Group6. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are close to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved. The Figure 4-39 shows the goodness of fit for the data with the statistics of Pearson’s and Spearman’s R.

Comparison | **Goodness of Fit** | Variance | Theory & Methods

Calculate

	Pearson	Spearman
X Axis:	0.998	0.997
Y Axis:	0.998	0.998

Figure 3-39 Pearson’s and Spearman’s R for the Goodness of Fit

To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two groups the two following networks for the two groups will appear. The Figure 4-40 shows the cognitive network for the Group1, the Figure 4-41 shows the Comparison Networks for Group1 and Group6 and the Figure 4-42 shows the cognitive network for the Group6.

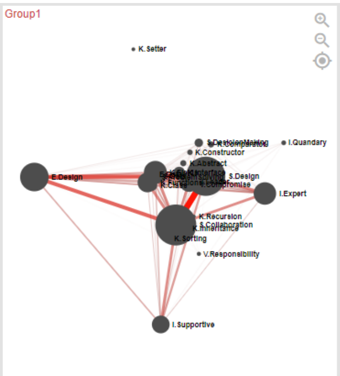


Figure 3-40 The Network of the centroid for Group1

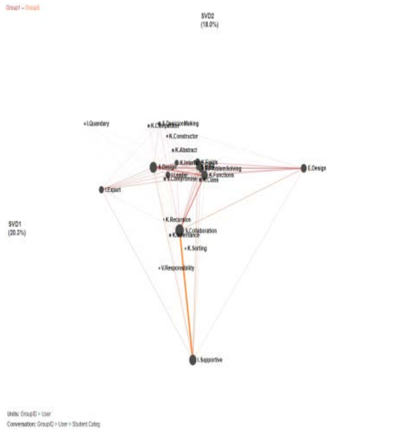


Figure 3-41 Comparison of Group1 and Group6

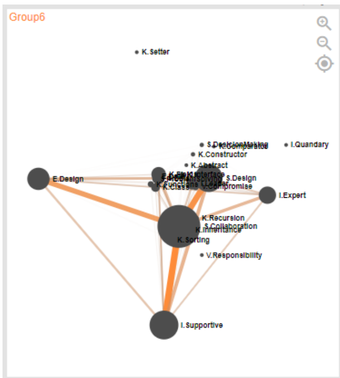


Figure 3-42 The Network of the centroid for Group6

There are many differences between the two groups analyzed. The differences worth presenting are:

The members of the Group 1 focused more on the Methods of the OOP because there are strong connections between Knowledge of Methods and Skills of Design (K.Methods - S.Design: 0.142) and Collaboration (K.Methods - S.Collaboration: 0.155).There are also stronger connections between the Skill of Design and the

Epistemology of Design (S.Design-E.Design: 0.098) and the Skill of Problem Solving (S.ProblemSolving-S.Design: 0.094) as well as between the Skill of Problem Solving and the Skill of Collaboration (S.ProblemSolving-S.Collaboration: 0.103).That means that the members had the good knowledge of the Methods needed to solve the exercise and the problem solving skills necessary for the solution.

The members of the Group 6 have stronger connections with the Skill of Collaboration and the identity of Supportive. Specifically there were strong connections between Skill of Collaboration and the Epistemology of Design (S.Collaboration - E.Design: 0.096) and the Skill of Data (S.Collaboration-S.Data: 0.107). The connections from the identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.282), the skill of Design (I.Supportive-S.Design: 0.103) and the Skill of Data (I.Supportive-S.Data: 0.092). That means that students of Group 1 have more advance knowledge and skills in comparison to the other group but the members of the Group 6 have better collaboration in the group with the identity of the Supportive to appear prominent which means that the members of the Group 6 helped more each other in order to solve the exercise.

3.3.2.6 Comparison Group1 and Group7

In this case there is not a significant difference on the first dimension but there is a significant difference on the second one:

- *Along the X axis, a two sample t test assuming unequal variance showed Group1 (mean=0.196, SD=0.884, N=3 was not statistically significantly different at the alpha=0.05 level from Group7 (mean=-0.039, SD=1.125, N=4; t(4.945)= 0.309, p=0.770, Cohen's d=0.227).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group1 (mean=2.228, SD=0.188, N=3 was statistically significantly different at the alpha=0.05 level from Group7 (mean=0.073, SD=0.466, N=4; t(4.155)= 8.385, p=0.001, Cohen's d=5.672).*

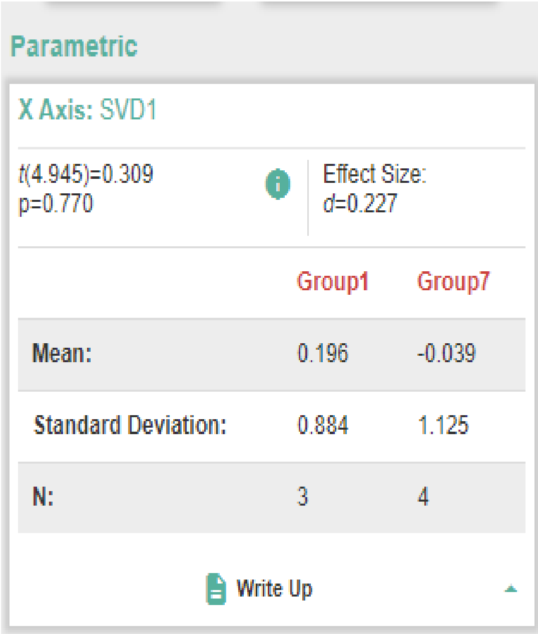


Figure 3-43 The results of the Independent T-Test for the first dimension

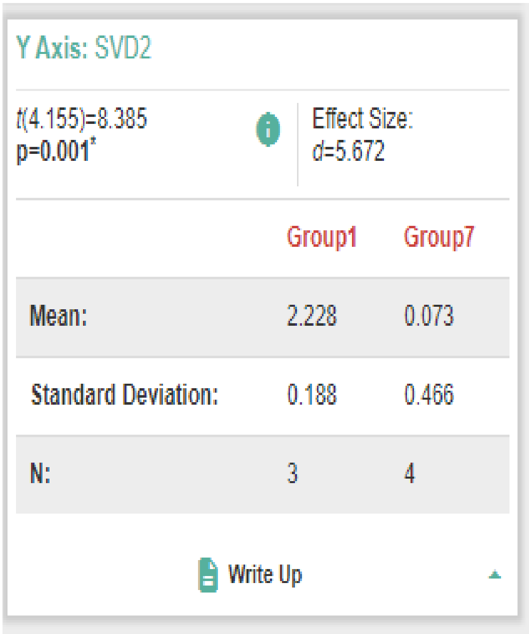


Figure 3-44 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-43 and Figure 4-44 show the independent t-test for the first and second dimension for Group1 and Group7. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are close to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved. The Figure 4-45 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R.

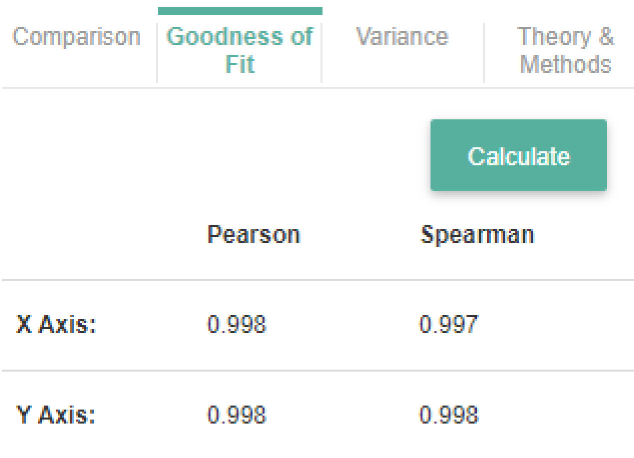


Figure 3-45 Pearson's and Spearman's R for the Goodness of Fit

To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two groups the two following networks for the two groups will appear. The Figure 4-46 shows the cognitive network for the Group1, the Figure 4-47 shows the Comparison Networks for Group1 and Group7 and the Figure 4-48 shows the cognitive network for the Group7.

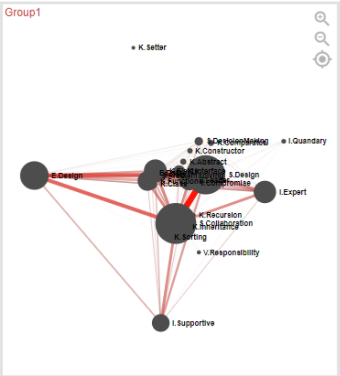


Figure 3-46 The Network of the centroid for Group1

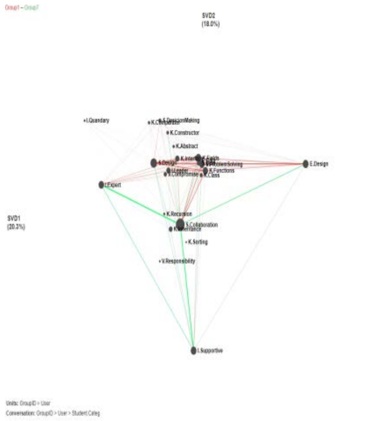


Figure 3-47 Comparison of Group1 and Group7

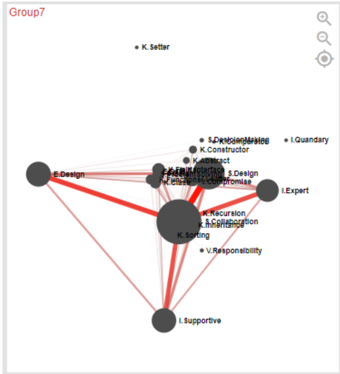


Figure 3-48 The Network of the centroid for Group7

There are many differences between the two groups analyzed. The differences worth presenting are:

The members of the Group 1 focused more on the Methods and the Fields of the data of the OOP because there are strong connections between Knowledge of Methods and the Skills of Design (K.Methods - S.Design: 0.092) and the Epistemology of Design (K.Methods - E.Design: 0.066). Also there are strong connections between the Knowledge of Fields and the Skills of Design (K.Fields - S.Design: 0.098) as well as the Epistemology of Design (K.Fields - E.Design: 0.073).There are also stronger connections between the Epistemology of Design and the Skill of Design (E.Design-S.Design: 0.086) as well as the Skill of Problem Solving (E.Design-S.ProblemSolving: 0.067). However there are strong connections between the Skill of Problem Solving and the Design (S.ProblemSolving-S.Design: 0.086). That means that the members had the knowledge and the skills needed.

The members of the Group7 have stronger connections with the Skill of Collaboration and the Identity of Expert. Specifically there were strong connections between Skill of Collaboration and the Epistemology of Design (S.Collaboration - E.Design: 0.088), the Skill of Data (S.Collaboration-S.Data: 0.065) and the Identity of Supportive (I.Supportive-S.Collaboration: 0.132). The connections from the Identity of Expert and the Identity of Supportive (I.Expert-I.Supportive: 0.082), and the Skill of Collaboration (I.Expert-S.Collaboration: 0.141) are also strong. That means that students ofGroup 1 have more advance knowledge and design skills in comparison to the other group but the members of the Group 7 have stronger collaboration and connections between the different identities of its members.

3.3.2.7 Comparison Group1 and Group8

In this case there is not a significant difference on the first dimension but there is a significant difference on the second one:

- *Along the X axis, a two sample t test assuming unequal variance showed Group1 (mean=0.196, SD=0.884, N=3 was not statistically significantly different at the alpha=0.05 level from Group8 (mean=0.272, SD=0.670, N=3; t(3.726)= -0.118, p=0.912, Cohen's d=0.097).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group1 (mean=2.228, SD=0.188, N=3 was statistically significantly different at the alpha=0.05 level from Group8 (mean=-0.320, SD=0.545, N=3; t(2.472)= 7.657, p=0.009, Cohen's d=6.252).*

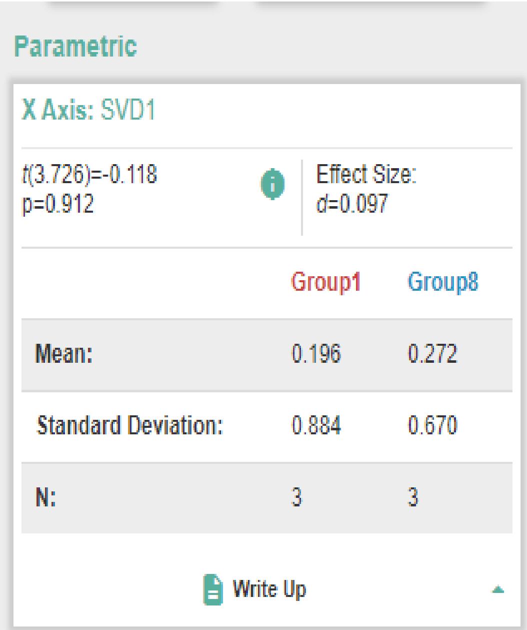


Figure 3-49 The results of the Independent T-Test for the first dimension

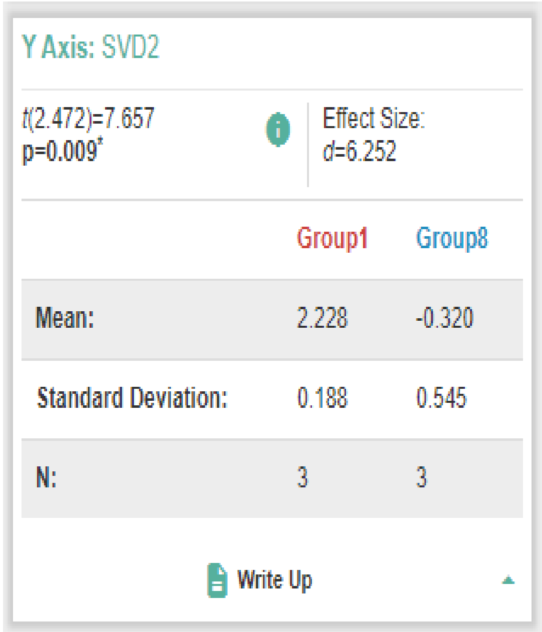
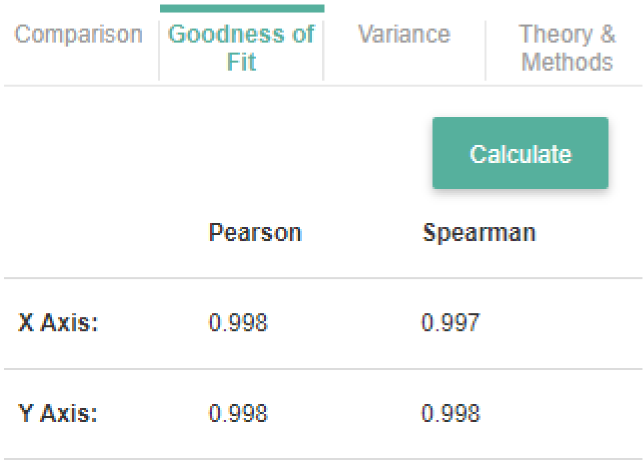


Figure 3-50 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-49 and Figure 4-50 show the independent t-test for the first and second dimension for Group1 and Group8. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson’s and Spearman’s r. In this case, both are close to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved. The Figure 4-51 shows the goodness of fit for the data with the statistics of Pearson’s and Spearman’s R.



	Pearson	Spearman
X Axis:	0.998	0.997
Y Axis:	0.998	0.998

Figure 3-51 Pearson's and Spearman's R for the Goodness of Fit

To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two groups the two following networks for the two groups will appear. The Figure 4-52 shows the cognitive network for the Group1, the Figure 4-53 shows the Comparison Networks for Group1 and Group8 and the Figure 4-54 shows the cognitive network for the Group8.

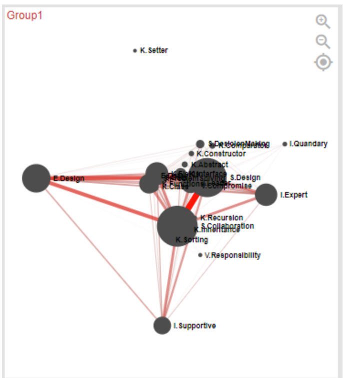


Figure 3-52 The Network of the centroid for Group1

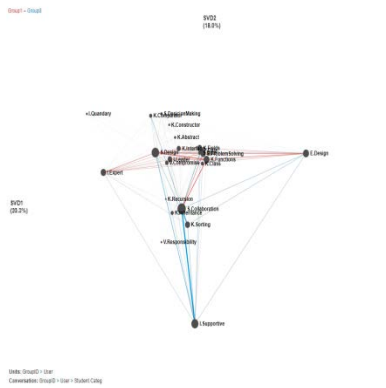


Figure 3-53 Comparison of Group1 and Group8

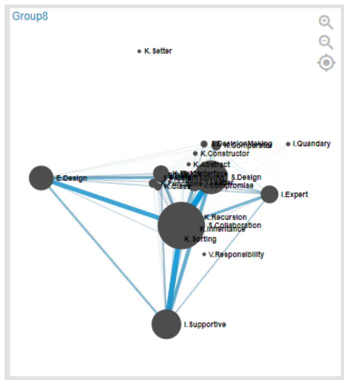


Figure 3-54 The Network of the centroid for Group8

There are many differences between the two groups analyzed. The differences worth presenting are:

The members of the Group 1 focused more on the Methods of the OOP because there are strong connections between Knowledge of Methods and the Skills of Design (K.Methods - S.Design: 0.117) and Collaboration (K.Methods - S.Collaboration: 0.121) as well as the Epistemology of Design (K.Methods - E.Design: 0.084).There are also stronger connections between the Skill of Problem Solving and the Skill of Design(S.Design-S.ProblemSolving: 0.084), and the Skill of Collaboration (S.Collaboration-S.ProblemSolving: 0.088).There were also a prominent connection between the Skill of Design and the Identity of Expert (I.Expert-S.Design: 0.078). That means that the members had the good knowledge and the skills needed to solve the exercise and collaborated over that.

The members of the Group 8 have stronger connections with the identity of Supportive and the Knowledge of Sorting. Specifically there were strong connections between Knowledge of Sorting and the Skill of Collaboration (K.Sorting - S.Collaboration: 0.141) and the Skill of Design (K.Sorting-S.Design: 0.071). The connections from the identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.182), the skill of Design (I.Supportive-S.Design: 0.097) are also evident. Lastly there is a strong connection between the Epistemology of Data and the Skill of Collaboration (S.Collaboration-E.Data: 0.08). That means that the Group 1 has better skills in comparison to the other group but the members of the Group 8focused more on the Sorting process in OOP .Lastly the identity of the Supportive appear prominent which means that the members of the Group 8 helped more each other in order to solve the exercise.

3.3.2.8 Comparison Group1 and Group9

In this case there is not a significant difference on the first dimension but there is a significant difference on the second one:

- *Along the X axis, a two sample t test assuming unequal variance showed Group1 (mean=0.196, SD=0.884, N=3 was not statistically significantly different at the alpha=0.05 level from Group9 (mean=-0.178, SD=1.838, N=3; t(2.879)= 0.318, p=0.772, Cohen's d=0.259).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group1 (mean=2.228, SD=0.188, N=3 was not statistically significantly different at the alpha=0.05 level from Group9 (mean=-0.796, SD=1.781, N=3; t(2.045)= 2.925, p=0.097, Cohen's d=2.388).*

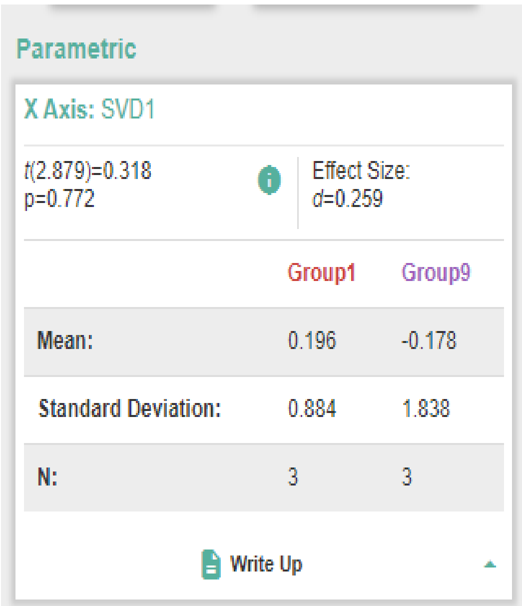


Figure 3-55 The results of the Independent T-Test for the first dimension

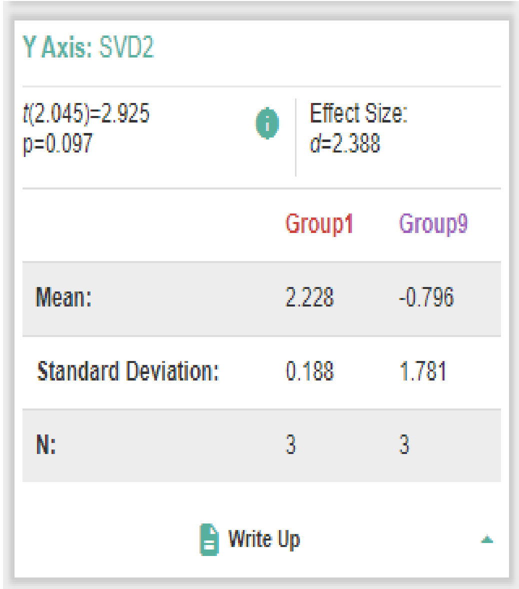


Figure 3-56 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-55 and Figure 4-56 show the independent t-test for the first and second dimension for Group1 and Group9. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are close to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved. The Figure 4-57 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R.

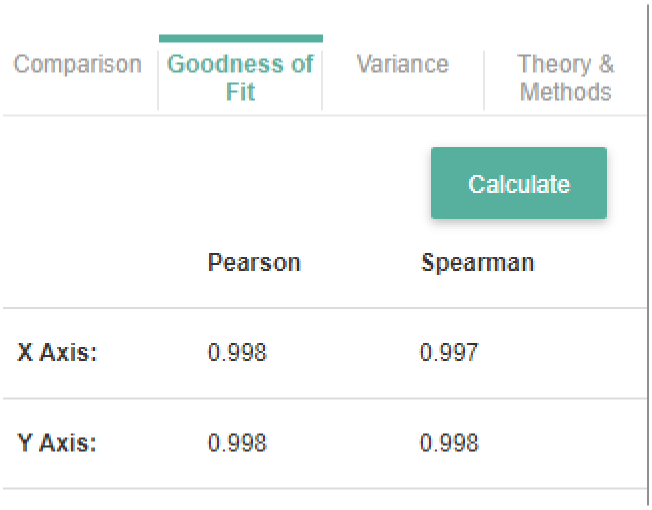


Figure 3-57 Pearson's and Spearman's R for the Goodness of Fit

To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two groups the two following networks for the two groups will appear. The Figure 4-58 shows the cognitive network for the Group1, the Figure 4-59 shows the Comparison Networks for Group1 and Group9 and the Figure 4-60 shows the cognitive network for the Group9.

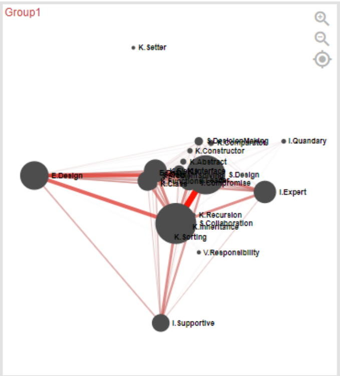


Figure 3-58 The Network of the centroid for Group1

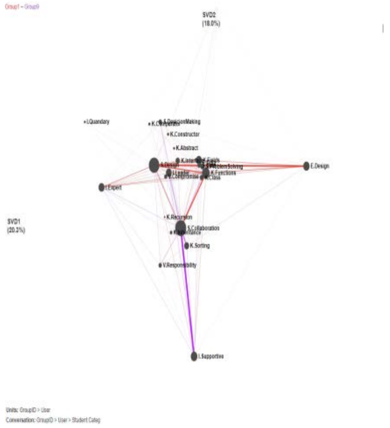


Figure 3-59 Comparison of Group1 and Group9

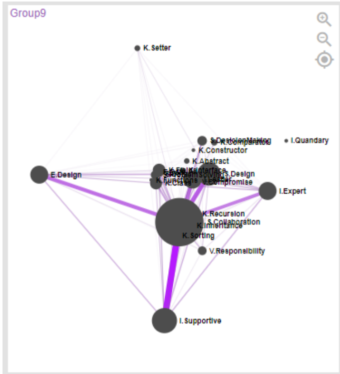


Figure 3-60 The Network of the centroid for Group9

There are many differences between the two groups analyzed. The differences worth presenting are:

The members of the Group 1 focused more on the Methods and the Fields of the data of the OOP. There are strong connections between the Knowledge of Methods and the Skills of Design (K.Methods - S.Design: 0.165), Collaboration (K.Methods - S.Collaboration: 0.197), the Epistemology of Design (K.Methods - E.Design: 0.104), the Identity of Expert (K.Functions-I.Expert: 0.09) and the Identity of Leader (K.Functions-I.Leader: 0.102). Also there are strong connections between K.Fields and the Skills of Design (K.Fields - S.Design: 0.113) as well as the Epistemology of Design (K.Fields - E.Design: 0.092).There are also stronger connections with the Skill of Design and the Epistemology of Design (S.Design-E.Design: 0.198) as well as the Identity of Expert(S.Design- I.Expert: 0.101). That means that the members had the good knowledge of the basics and the skills needed to solve the exercise and collaborated over that.

The members of the Group 9 have stronger connections with the Skill of Collaboration. Specifically there were strong connections between Skill of Collaboration and the Knowledge of Class in Java (K.Class-S.Collaboration: 0.099) and the Knowledge of Sorting (K.Sorting-S.Collaboration: 0.147). The connections from the Skill of Collaboration and the identity of Leader (I.Leader-S.Collaboration:0.094) and the Identity of Supportive (I.Supportive-S.Collaboration: 0.269) are also evident. Lastly there are strong connections between the Skill of Collaboration and the Value of

Responsibility (S.Collaboration-V.Responsibility: 0.075) and Skill of Decision Making (S.Collaboration-S.DesicionMaking: 0.06). That means that students of Group 1 have better developing skills in comparison to the other group but the members of the Group 9 have prominent the identity of the Supportive and they have better collaboration with each other.

3.3.2.9 Comparison Group1 and Group 10

In this case there is not a significant difference on the first dimension but there is a significant difference on the second one:

- *Along the X axis, a two sample t test assuming unequal variance showed Group1 (mean=0.196, SD=0.884, N=3 was not statistically significantly different at the alpha=0.05 level from Group10 (mean=-0.824, SD=2.157, N=4; t(4.180)= 0.856, p=0.439, Cohen's d=0.579).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group1 (mean=2.228, SD=0.188, N=3 was statistically significantly different at the alpha=0.05 level from Group10 (mean=-1.450, SD=0.636, N=4; t(3.668)= 10.949, p=0.001, Cohen's d=7.260).*

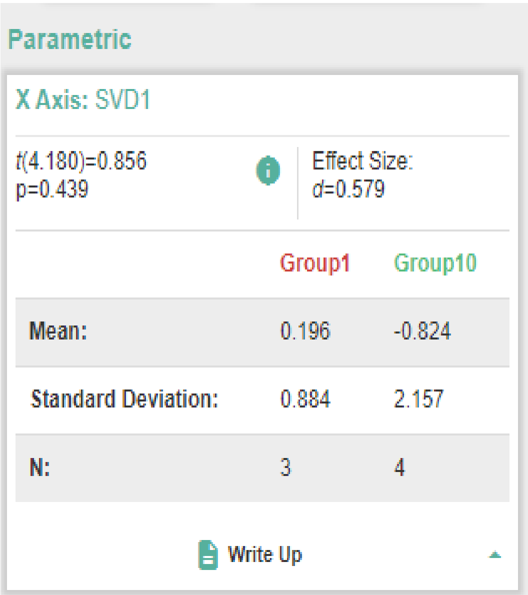


Figure 3-61 The results of the Independent T-Test for the first dimension

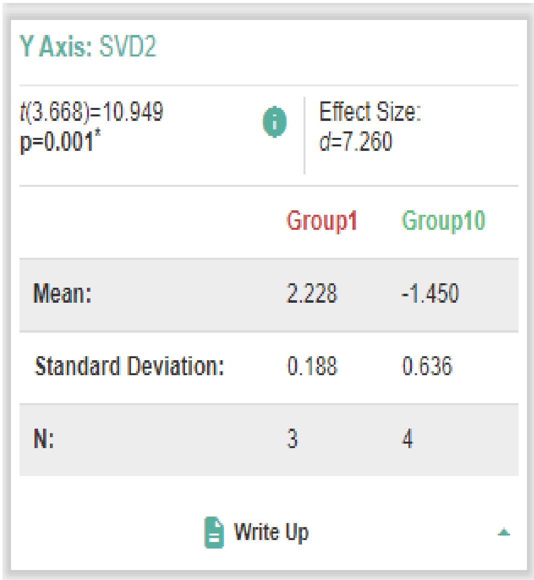


Figure 3-62 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-61 and Figure 4-62 show the independent t-test for the first and second dimension for Group1 and Group10. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are close to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved. The Figure 4-63 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R.

Comparison | **Goodness of Fit** | Variance | Theory & Methods

Calculate

	Pearson	Spearman
X Axis:	0.998	0.997
Y Axis:	0.998	0.998

Figure 3-63 Pearson's and Spearman's R for the Goodness of Fit

To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two groups the two following networks for the two groups will appear. The Figure 4-64 shows the cognitive network for the Group1, the Figure 4-65 shows the Comparison Networks for Group1 and Group10 and the Figure 4-66 shows the cognitive network for the Group10.

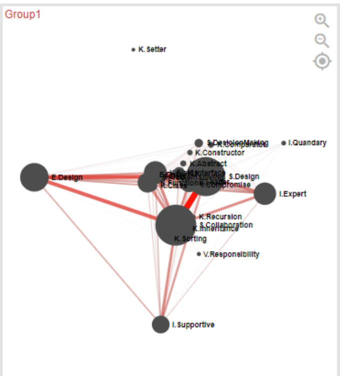


Figure 3-64 The Network of the centroid for Group1

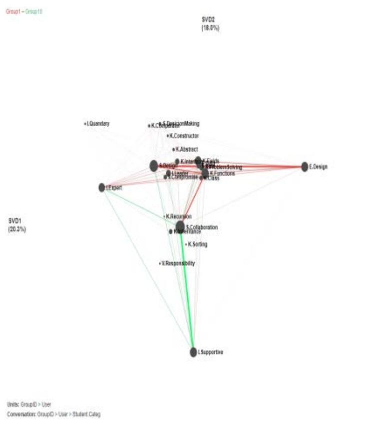


Figure 3-65 Comparison of Group1 and Group10

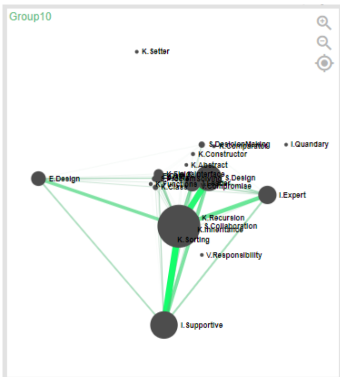


Figure 3-66 The Network of the centroid for Group10

There are many differences between the two groups analyzed. The differences worth presenting are:

The members of the Group 1 focused more on the Methods of the OOP because there are strong connections between Knowledge of Methods and the Skill of Design (K.Methods - S.Design: 0.160) and Collaboration (K.Methods - S.Collaboration: 0.193), and the Epistemology of Design (K.Functions- E.Design: 0.100). There are also stronger connections between the Skill of Design and the epistemology of Design (S.Design-

E.Design: 0.205) as well as the Knowledge of Fields (S.Design-K.Fields:0.11). That means that the members had the good knowledge of the methods needed and the Problem Solving skill to solve the exercise and collaborated over that.

The members of the Group 10 have stronger connections with the Skill of Collaboration and the identity of Supportive. Specifically there were strong connections between Skill of Collaboration and the identities of Supportive (I.Supportive-S.Collaboration: 0.286) and Expert (I.Expert-S.Collaboration: 0.107) as well as the Skill of Data (S.Collaboration-S.Data: 0.079). The connections from the identity of Supportive and the skill of Design (I.Supportive-S.Design: 0.098) and the identity of Expert (I.Supportive-I.Expert: 0.085). That means that students of Group 1 have more advance knowledge and skills in comparison to the other group but the members of the Group 10 have distinctive roles and better collaboration. Also the identity of Supportive appears prominent which means that the members of the Group 10 helped more each other in order to solve the exercise.

3.3.3 RQ2a. and RQ2b.

RQ2a. What types of connections between codes are made by groups in the High-to-High category? What types of connections between codes are made by groups in the High/Low-to-High category? What types of connections between codes are made by groups in the High/Low-to-Low category?

RQ2b. Is there a significant difference between the discourse networks of groups of the three categories: High-to-High, High/Low-to-High and High/Low-to-Low?

In this case we examine the ability of students to solve an Object Oriented assignment using Java based the groups that they created by themselves.

The units, the codes and the conversation that were chosen are the following:

- Units: Category2,GroupID, UserID
- Conversation: GroupID,Category2, UserID
- Stanza Window: Moving Window Consisted of 4 Lines
- Codes: All the codes given in the file
- Comparison: Category2

The figure we see in ENA WebKit, after selecting the parameters, is the following (Fig.67), where we can see the centroids and the confidence intervals for the three Group Categories.

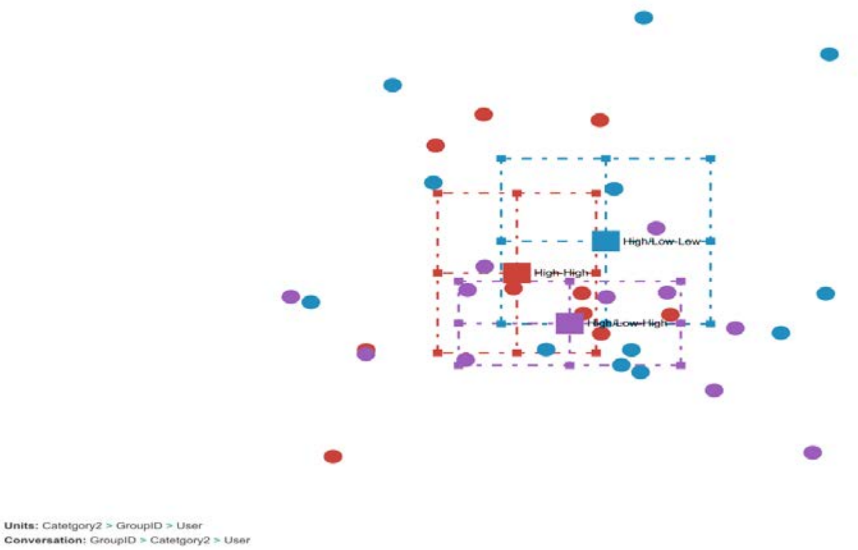


Figure 3-67 The Centroids of the Three Categories Groups (Red: High-to-High, Purple: High/Low-to-High and Blue: High/Low-to-Low)

By default, the ENA space in which the centroids are displayed is determined by the first (x) and second (y) dimensions, the dimensions that represent the largest data variation. The numbers in parentheses next to the axis labels indicate the percentage of variation in the data relating to these dimensions. In this case, the dimension x (SVD1) represents 18.7% of the variance in the data and the dimension y (SVD2) represents 17.4%.

As shown in Fig. 87, the three Group Categories are slightly different. To determine the difference more accurately, we can perform an independent samples t - test. To do this, simply select the two samples we want to compare from the drop-down menu on the left, in the tab "Stats". When we do this, we will see averages for the two samples, along with the t-score, the p value and Cohen's d, a measure of the magnitude of the effect. In order to analyze them further we will present the differences of the connections between the two groups. This difference is calculated by subtracting the weight of the edge of one network from the other. The numbers in the brackets is the numeric differences of each edge between the two Groups or Categories.

3.3.3.1 Correlation

The strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson’s and Spearman’s r. In this case, both are equal to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved. The Figure 4-68 shows the goodness of fit for the data with the statistics of Pearson’s and Spearman’s R.

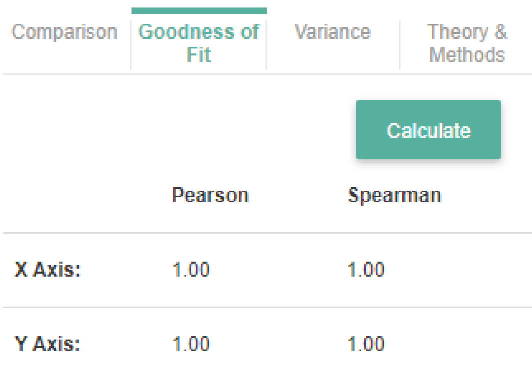


Figure 3-68 Correlation for the 3 Categories

3.3.3.2 Comparison High-to-High and High/Low-to-High

In this case, the difference on the first and the second dimension is not significantly different:

- *Along the X axis, a two sample t test assuming unequal variance showed High-High (mean=-0.53, SD=1.17, N=10 was not statistically significantly different at the alpha=0.05 level from High/Low-High (mean=0.03, SD=1.74, N=11; t(17.56)= 0.87, p=0.40, Cohen's d=0.37).*
- *Along the Y axis, a two sample t test assuming unequal variance showed High-High (mean=0.07, SD=1.57, N=10 was not statistically significantly different at the alpha=0.05 level from High/Low-High (mean=-0.64, SD=0.88, N=11; t(13.87)= -1.26, p=0.23, Cohen's d=0.56).*

X Axis: SVD1

t(17.56)=0.87
p=0.40

Effect Size:
d=0.37

	High-High	High/Low-High
Mean:	-0.53	0.03
Standard Deviation:	1.17	1.74
N:	10	11

Write Up

Figure 3-69 The results of the Independent T-Test for the first dimension

Y Axis: SVD2

t(13.87)=-1.26
p=0.23

Effect Size:
d=0.56

	High-High	High/Low-High
Mean:	0.07	-0.64
Standard Deviation:	1.57	0.88
N:	10	11

Write Up

Figure 3-70 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-69 and Figure 4-70 show the independent t-test for the first and second dimension for the categories High-to-High and High/Low-to-High. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two categories the two following networks for the two categories will appear. The Figure 4-71 shows the cognitive network for the category High-to-High, the Figure 4-72 shows the Comparison Networks for the categories High-to-High and High/Low-to-High and the Figure 4-73 shows the cognitive network for the category High/Low-to-High.

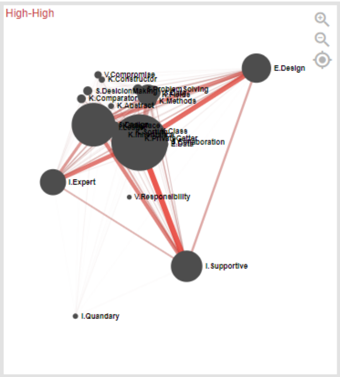


Figure 3-71 Network of groups in the High-to-High category

Figure 3-72 Comparison of High-to-High and High/Low-to-High group categories

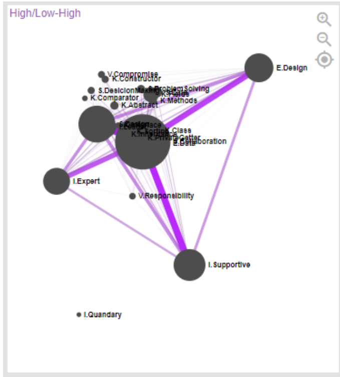


Figure 3-73 Network of groups in the High/Low-to-High category

There are no major differences between the two categories analyzed. The differences worth presenting are:

The members of the category *High-to-High* have feature more connections between the Skill and the Knowledge nodes than the other category. Specifically there are stronger connections between the Knowledge of Methods and Fields, and the Skills of Collaboration (K.Methods-S.Collaboration: 0.039) (*K.Fields- S.Collaboration: 0.048)* and Design (K.Methods-S.Design: 0.035) (*K.Fields- S.Design: 0.050*). The Skill of Collaboration has strong connections with the Epistemology of Data (*S.Collaboration-E.Data: 0.033)*, the Knowledge of Sorting (K.Sorting- S.Collaboration: 0.033) and of the Comparator (*K.Comparator- S.Collaboration: 0.031*) and the Skill of Design (S.Design-S.Collaboration: 0.040). There are also strong connections between the identity of the Leader and the Epistemology code of Design (I.Leader-E.Design: 0.033), the Knowledge Code of Methods (K.Methods- I.Leader: 0.032) and the Skills of Design *(S.Design-I.Leader: 0.040)* and Collaboration (I.Leader-S.Collaboration: 0.045). This means that the identity of the Leader is more prominent in this category.

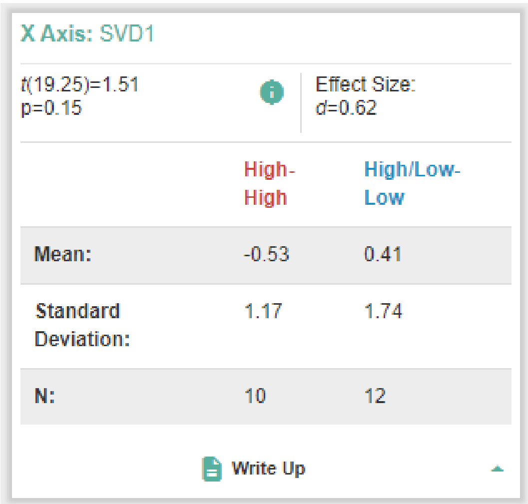
The members of the category *High/Low-to-High* have more connections with the Skill of Collaboration and the identity of Supportive. Specifically there were strong connections between Skill of Collaboration and the identities of Supportive (I.Supportive-S.Collaboration: 0.059) and Expert (I.Expert-S.Collaboration: 0.046) as well as the Epistemology of Design (S.Collaboration-E.Design: 0.059). The connections from the identity of Supportive and the Epistemology of Design (I.Supportive-E.Design: 0.026) and the identity of Expert (I.Supportive-I.Expert: 0.031). The Knowledge node more prominent in this category is the Class and its connections with the Skills of Design

(K.Class-S.Design: 0.034) and Collaboration (K.Class-S.Collaboration: 0.046). That means that category High-to-High has more advance knowledge and skills in comparison to the other group but the members of the Category *High/Low-to-High* have better collaboration and its members appear to help more each other in order to solve the exercise.

3.3.3.3 Comparison High-to-High and High/Low-to-Low

In this case, the difference on the first and the second dimension is not significantly different:

- *Along the X axis, a two sample t test assuming unequal variance showed High-High (mean=-0.53, SD=1.17, N=10 was not statistically significantly different at the alpha=0.05 level from High/Low-to-Low (mean=0.41, SD=1.74, N=12; t(19.25)= 1.51, p=0.15, Cohen's d=0.62).*
- *Along the Y axis, a two sample t test assuming unequal variance showed High-High (mean=0.07, SD=1.57, N=10 was not statistically significantly different at the alpha=0.05 level from High/Low-to-Low (mean=0.52, SD=1.84, N=12; t(19.97)= 0.62, p=0.54, Cohen's d=0.26).*



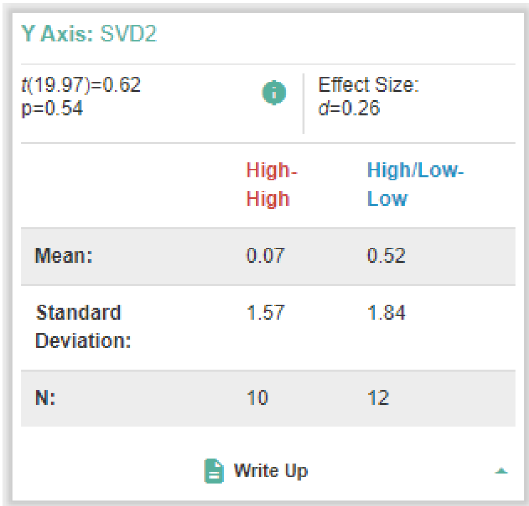
X Axis: SVD1

t(19.25)=1.51 p=0.15 | Effect Size: d=0.62

	High-High	High/Low-Low
Mean:	-0.53	0.41
Standard Deviation:	1.17	1.74
N:	10	12

Write Up

Figure 3-74 The results of the Independent T-Test for the first dimension



Y Axis: SVD2

t(19.97)=0.62 p=0.54 | Effect Size: d=0.26

	High-High	High/Low-Low
Mean:	0.07	0.52
Standard Deviation:	1.57	1.84
N:	10	12

Write Up

Figure 3-75 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-74 and Figure 4-75 show the independent t-test for the first and second dimension for the categories High-to-High and High/Low-to-Low. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two categories the two following networks for the two categories will appear. The Figure 4-76 shows the cognitive network for the category High-to-High, the Figure 4-77 shows the Comparison Networks for the categories High-to-High and High/Low-to-Low and the Figure 4-78 shows the cognitive network for the category High/Low-to-Low.

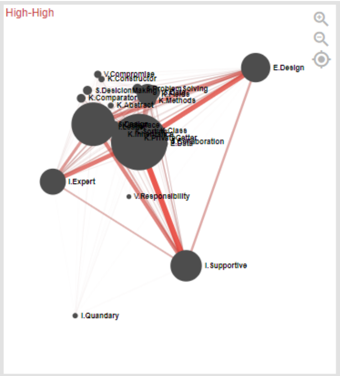


Figure 3-76 Network of groups in the High-to-High category

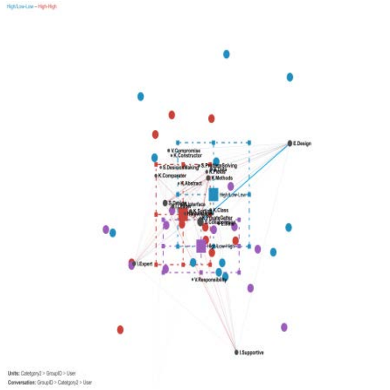


Figure 3-77 Comparison of High-to-High and High/Low-to-Low

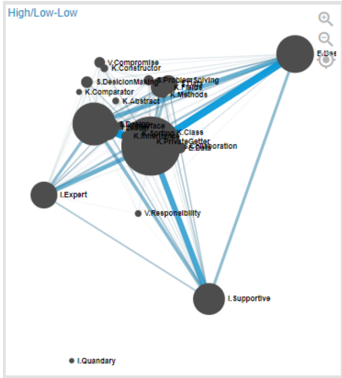


Figure 3-78 Network of groups in the High/Low-to-Low category

There are no major differences between the two categories analyzed. The differences worth presenting are:

The members of the category *High-to-High* have feature more connections between the Skill and the Knowledge nodes than the other category. Specifically there are stronger connections between the Knowledge of Methods and Fields, and the Skills of Collaboration (K.Methods-S.Collaboration: 0.069) and Design (K.Methods-S.Design: 0.055) and between the Knowledge of Fields and the Skill of Design (*K.Fields-S.Design: 0.044*). The Skill of Collaboration has strong connections with the Knowledge of Sorting (K.Sorting- S.Collaboration: 0.057) and of the Comparator (*K.Comparator-S.Collaboration: 0.036*) and the Skill of Design (S.Design-S.Collaboration: 0.039). There are also strong connections between the identity of Supportive and the Skill of Design (I.Supportive-S.Design: 0.046) and between the Leader and the Skill of Collaboration (I.Leader-S.Collaboration: 0.039). This means that this category has more connections between the Knowledge and Skill Codes.

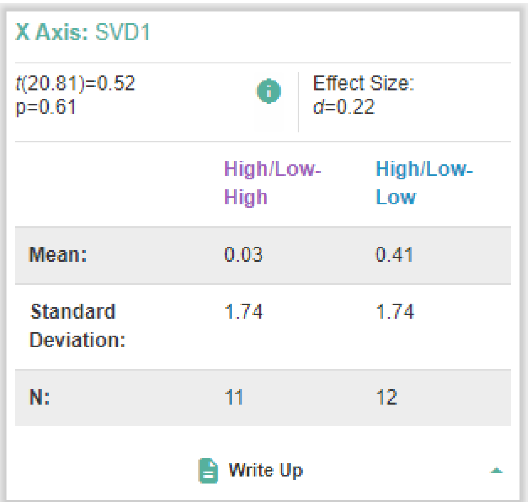
The members of the category *High/Low-to-Low* have more connections with the Epistemology of Design. Specifically there were strong connections between Epistemology of Design and Skill of Collaboration (E.Design-S.Collaboration: 0.102), the identity of Expert (I.Expert-E.Design: 0.036)as well as the Skill of Data (S.Data-E.Design: 0.032). The connections between the Skill of Collaboration and the Knowledge of Class (K.Class- S.Collaboration: 0.041) and the Skill of Data (S.Data-

S.Collaboration: 0.035) are also strong. In general this category forms weaker connection in comparison to the *High-to-High category and its members tend to confirm more often the ideas of others than propose new ones.* That means that category High-to-High has more advance knowledge and skills in comparison to the other group but the members of the Category *High/Low-to-Low* collaborate and confirm each other more times in order to solve the exercise.

3.3.3.4 Comparison High/Low-to-High and High/Low-to-Low

In this case, the difference on the first and the second dimension is not significantly different:

- *Along the X axis, a two sample t test assuming unequal variance showed High/Low-High (mean=0.03, SD=1.74, N=11 was not statistically significantly different at the alpha=0.05 level from High/Low-to-Low (mean=0.41, SD=1.74, N=12; t(20.81)= 0.52, p=0.61, Cohen's d=0.22).*
- *Along the Y axis, a two sample t test assuming unequal variance showed High/Low-High (mean=-0.64, SD=0.88, N=11 was not statistically significantly different at the alpha=0.05 level from High/Low-to-Low (mean=0.52, SD=1.84, N=12; t(16.13)= 1.95, p=0.07, Cohen's d=0.79).*



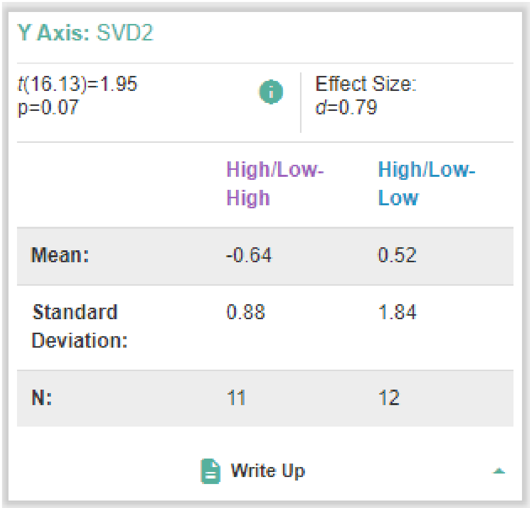
X Axis: SVD1

t(20.81)=0.52 p=0.61 | Effect Size: d=0.22

	High/Low-High	High/Low-Low
Mean:	0.03	0.41
Standard Deviation:	1.74	1.74
N:	11	12

Write Up

Figure 3-79 The results of the Independent T-Test for the first dimension



Y Axis: SVD2

t(16.13)=1.95 p=0.07 | Effect Size: d=0.79

	High/Low-High	High/Low-Low
Mean:	-0.64	0.52
Standard Deviation:	0.88	1.84
N:	11	12

Write Up

Figure 3-80 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-79 and Figure 4-80 show the independent t-test for the first and second dimension for the categories High/Low-to-High and High/Low-to-Low. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two categories the two following networks for the two categories will appear. The Figure 4-81 shows the cognitive network for the category High/Low-to-High, the Figure 4-82 shows the Comparison Networks for the categories High/Low-to-High and High/Low-to-Low and the Figure 4-83 shows the cognitive network for the category High/Low-to-Low.

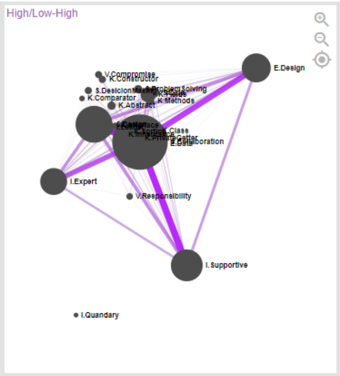


Figure 3-81 Network of groups in the High/Low-to-High category

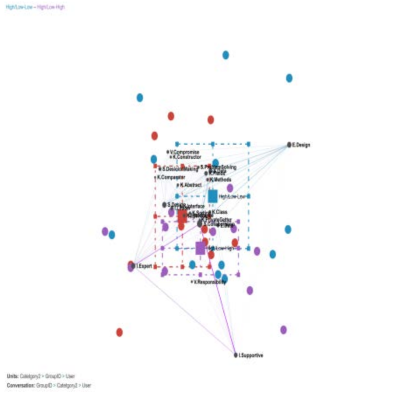


Figure 3-82 Comparison of High/Low-to-High and High/Low-to-Low

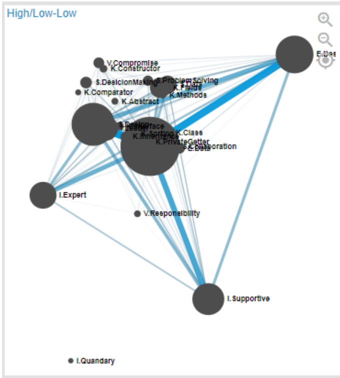


Figure 3-83 Network of groups in the High/Low-to-Low category

There are no major differences between the two categories analyzed. The differences worth presenting are:

The members of the category *High/Low-to-High* have more connections with the Skill of Collaboration and the identity of Supportive. Specifically there were strong connections between Skill of Collaboration and the identities of Supportive (I.Supportive-S.Collaboration: 0.085) and Expert (I.Expert-S.Collaboration: 0.065). The connections between the identity of Supportive and identity of Expert (I.Supportive-I.Expert: 0.045).That means that category *High/Low-to-High* have better collaboration and its members appear to help more each other in order to solve the exercise because the role of Supportive is more prominent in this category.

The members of the category *High/Low-to-Low* have stronger connections majority of the codes but the difference between them is not larger than 0.05/1. Specifically there were strong connections between Epistemology of Design and Skill of Collaboration (E.Design-S.Collaboration: 0.044), the identity of Leader (I.Leader-E.Design: 0.032) as well as the Skill of Data (S.Data-E.Design: 0.042) and Design (S.Design-E.Design: 0.042). The connections between the Skill of Data and the Knowledge of Fields (K.Fields- S.Collaboration: 0.038) and the Skill of Collaboration (S.Data- S.Collaboration: 0.035) and Design (S.Design- S.Data: 0.034) are also strong. In general this category forms more connection in comparison to the *High/Low-to-High* category and its members tend to confirm more often the ideas of others than propose new ones.

3.3.4 RQ3. *Is there a significant difference between the discourse networks of groups of the same Category?*

In this case we examine the ability of the students to solve an Object Oriented assignment using Java based the groups that they created by themselves.

The units, the codes and the conversation that were chosen are the following:

- Units: GroupID, UserID
- Conversation: GroupID, UserID,Student.Categ
- Stanza Window: Moving Window Consisted of 4 Lines
- Codes: All the codes given in the file
- Comparison: GroupID

3.3.4.1 High-to-High Groups' Category

The figure we see in ENA WebKit, after selecting the parameters, is the following (Figure 4-84), where we can see the centroids and the confidence intervals for the three groups of High-to-High category.

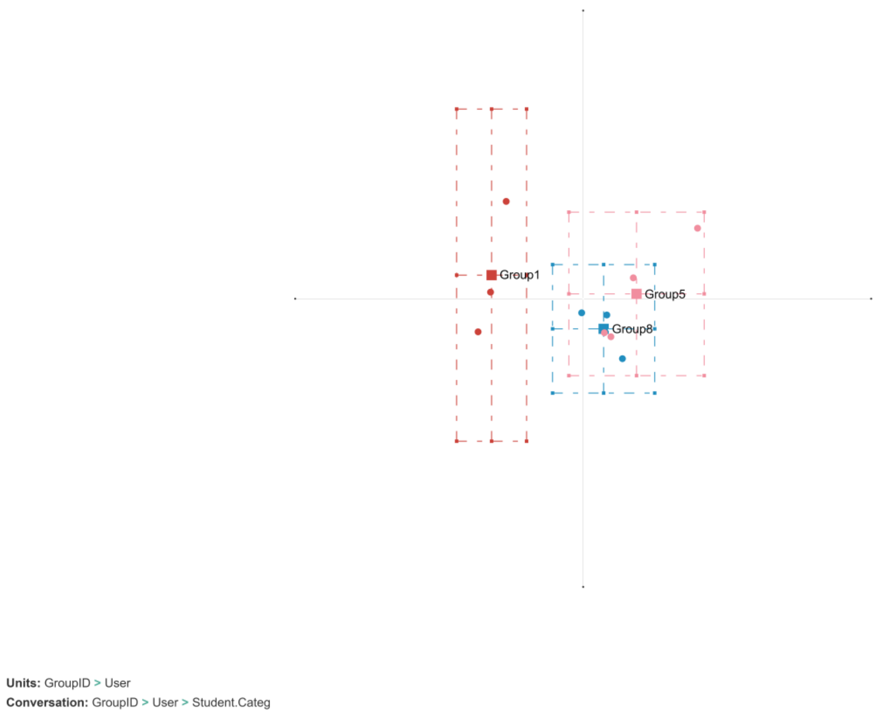


Figure 3-84 The Centroids networks of the Three Groups of the category High-to-High (RED: Group1, PINK: Group5, BLUE: Group8)

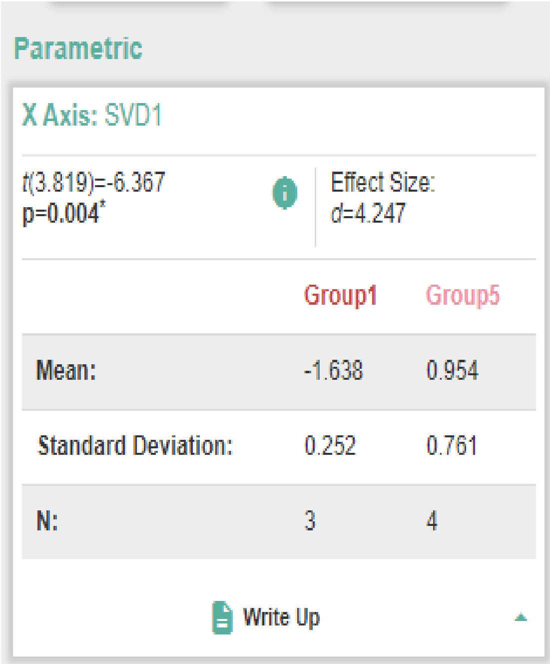
By default, the ENA space in which the centroids are displayed is determined by the first (x) and second (y) dimensions, the dimensions that represent the largest data variation. The numbers in parentheses next to the axis labels indicate the percentage of variation in the data relating to these dimensions. In this case, the dimension x (SVD1) represents 39.1% of the variance in the data and the dimension y (SVD2) represents 19.9%.

As shown in figure, the three Groups of category High-to-High are slightly different. To determine the difference more accurately, we can perform an independent samples t - test. To do this, simply select the two samples we want to compare from the drop-down menu on the left, in the tab "Stats". When we do this, we will see averages for the two samples, along with the t-score, the p value and Cohen's d, a measure of the magnitude of the effect. In order to analyze them further we will present the differences of the connections between the two groups. This difference is calculated by subtracting the weight of the edge of one network from the other. The numbers in the brackets is the numeric differences of each edge between the two Groups or Categories.

Comparison Group1 and Group5

In this case there is not a significant difference on the second dimension but there is a significant difference on the first one:

- *Along the X axis, a two sample t test assuming unequal variance showed Group1 (mean=-1.638, SD=0.252, N=3 was statistically significantly different at the alpha=0.05 level from Group5 (mean=0.954, SD=0.761, N=4; t(3.819)= -6.367, p=0.004, Cohen's d=4.247).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group1 (mean=0.422, SD=1.195, N=3 was not statistically significantly different at the alpha=0.05 level from Group5 (mean=0.087, SD=0.918, N=4; t(3.682)= 0.405, p=0.708, Cohen's d=0.323).*



Parametric

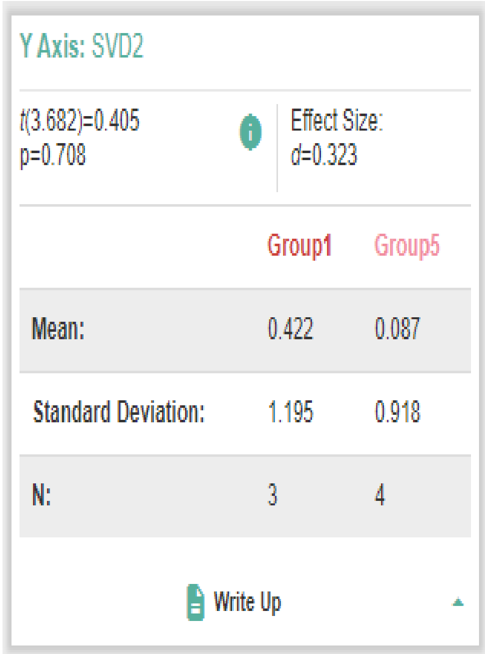
X Axis: SVD1

$t(3.819)=-6.367$, $p=0.004^*$ — Effect Size: $d=4.247$

	Group1	Group5
Mean:	-1.638	0.954
Standard Deviation:	0.252	0.761
N:	3	4

Write Up

Figure 3-85 The results of the Independent T-Test for the first dimension



Y Axis: SVD2

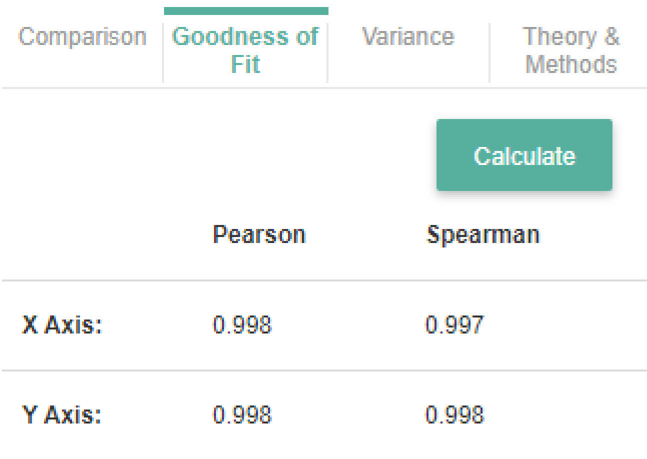
$t(3.682)=0.405$, $p=0.708$ — Effect Size: $d=0.323$

	Group1	Group5
Mean:	0.422	0.087
Standard Deviation:	1.195	0.918
N:	3	4

Write Up

Figure 3-86 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-85 and Figure 4-86 show the independent t-test for the first and second dimension for Group1 and Group5. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are close to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved.



Comparison | Goodness of Fit | Variance | Theory & Methods

Calculate

	Pearson	Spearman
X Axis:	0.998	0.997
Y Axis:	0.998	0.998

Figure 3-87 Pearson's and Spearman's R for the Goodness of Fit

The Figure 4-87 shows the goodness of fit for the data with the statistics of Pearson’s and Spearman’s R. To determine the differences in link structures between the two groups we look at the equiload projections. Thus clicking on the centroids of these two groups the two following networks for the two groups will appear. The Figure 4-88 shows the cognitive network for the Group1, the Figure 4-89 shows the Comparison Networks for Group1 and Group5 and the Figure 4-90 shows the cognitive network for the Group5.

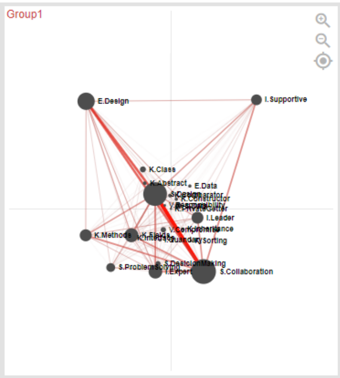


Figure 3-88 The Network of the centroid for Group1

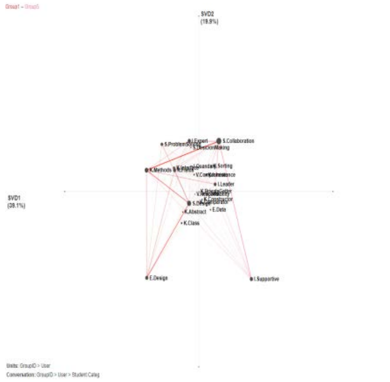


Figure 3-89 Comparison of Group1 and Group5

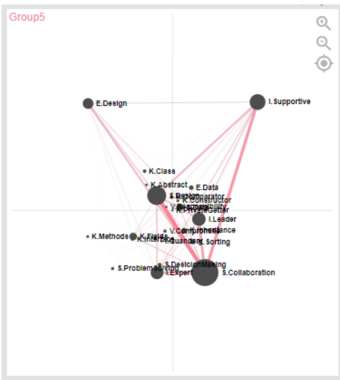


Figure 3-90 The Network of the centroid for Group5

There are many differences between the two groups analyzed. The differences worth presenting are:

The members of the Group 1 focused more on the Methods of the OOP because there are strong connections between the Knowledge of Methods and the Skills of Design (K.Methods - S.Design: 0.155) and Collaboration (K.Methods - S.Collaboration: 0.185), and the Identity of Leader (K.Methods-I.Leader: 0.098). There are also stronger connections between the Skill of Design and the epistemology of Design (S.Design-E.Design: 0.157) as well as between the skill of Collaboration and the skill of Problem Solving(S.ProblemSolving-S.Collaboration: 0.091) and the Skill of Design (S.ProblemSolving-S.Design: 0.081).That means that the members had the good knowledge of the methods needed and the Problem Solving skill to solve the exercise and collaborated over that.

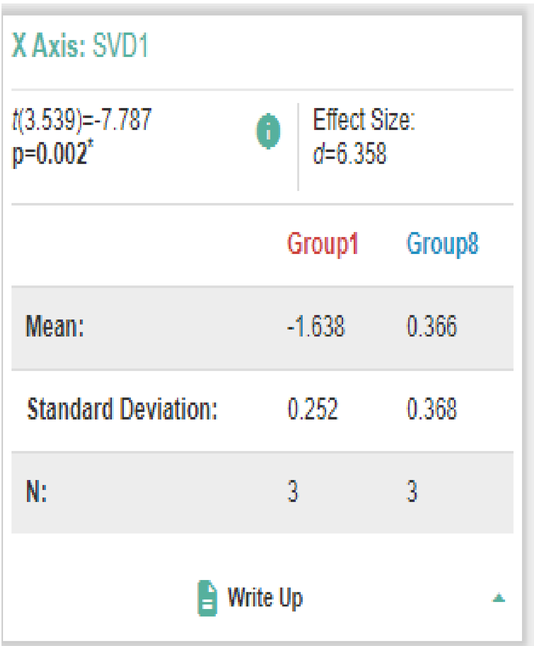
The members of the Group5 have stronger connections with the Skill of Collaboration and the identity of Supportive. Specifically there were strong connections between Skill of Collaboration and the Identities of Supportive (I.Supportive-S.Collaboration: 0.173), Leader (I.Leader-S.Collaboration: 0.138) and Expert (I.Expert-

S.Collaboration: 0.111). The connections from the identity of Supportive and the skill of Design (I.Supportive-S.Design: 0.089) and the identity of Leader (I.Supportive-I.Leader: 0.09). That means that the Group 1 has more advance knowledge and skills in comparison to the other group but the members of the Group 5 have distinctive roles and collaboration between the members of the group with the identity of the Supportive to appear prominent which means that the members of the Group 5 helped more each other in order to solve the exercise.

Comparison Group1 and Group8

In this case there is not a significant difference on the second dimension but there is a significant difference on the first one:

- *Along the X axis, a two sample t test assuming unequal variance showed Group1 (mean=-1.638, SD=0.252, N=3 was statistically significantly different at the alpha=0.05 level from Group8 (mean=0.366, SD=0.368, N=3; t(3.539)= -7.787, p=0.002, Cohen's d=6.358).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group1 (mean=0.422, SD=1.195, N=3 was not statistically significantly different at the alpha=0.05 level from Group8 (mean=-0.538, SD=0.462, N=3; t(2.585)= 1.298, p=0.298, Cohen's d=1.060).*



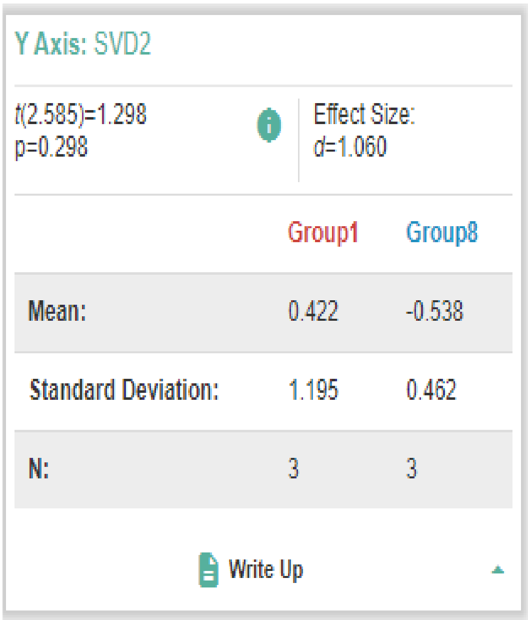
X Axis: SVD1

t(3.539)=-7.787 p=0.002* — Effect Size: d=6.358

	Group1	Group8
Mean:	-1.638	0.366
Standard Deviation:	0.252	0.368
N:	3	3

Write Up

Figure 3-91 The results of the Independent T-Test for the first dimension



Y Axis: SVD2

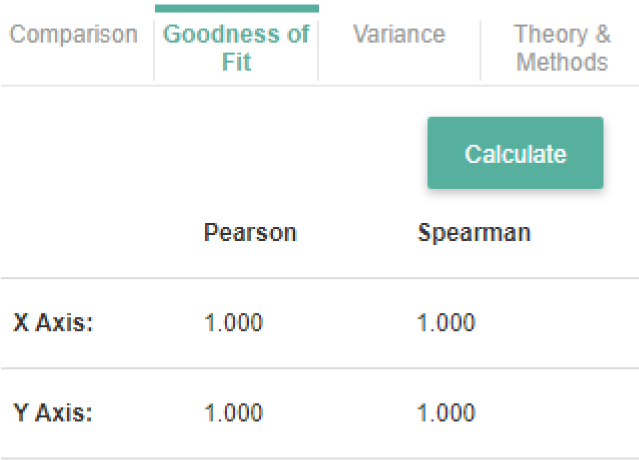
t(2.585)=1.298 p=0.298 — Effect Size: d=1.060

	Group1	Group8
Mean:	0.422	-0.538
Standard Deviation:	1.195	0.462
N:	3	3

Write Up

Figure 3-92 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-91 and Figure 4-92 show the independent t-test for the first and second dimension for Group1 and Group8. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are equal to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved.



	Pearson	Spearman
X Axis:	1.000	1.000
Y Axis:	1.000	1.000

Figure 3-93 Pearson's and Spearman's R for the Goodness of Fit

The Figure 4-93 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two groups the two following networks for the two groups will appear. The Figure 4-94 shows the cognitive network for the Group1, the Figure 4-95 shows the Comparison Networks for Group1 and Group8 and the Figure 4-96 shows the cognitive network for the Group8.

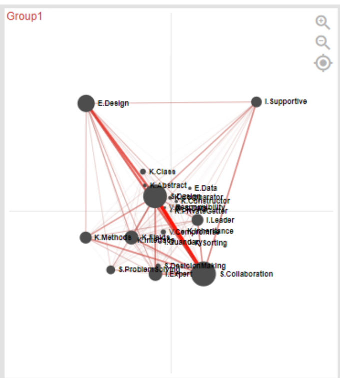


Figure 3-94 The Network of the centroid for Group1

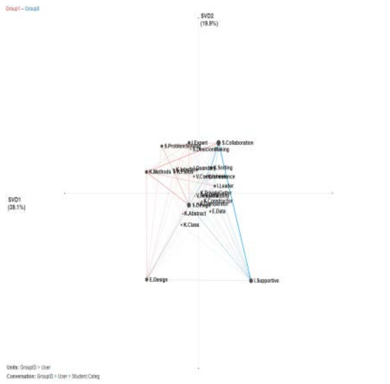


Figure 3-95 Comparison of Group1 and Group8

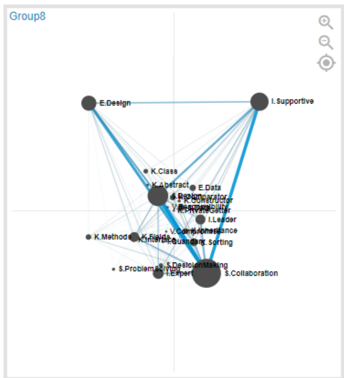


Figure 3-96 The Network of the centroid for Group8

There are many differences between the two groups analyzed. The differences worth presenting are:

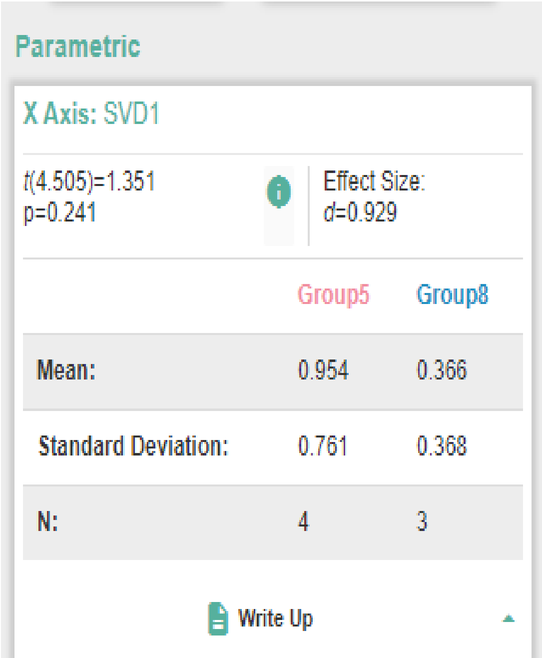
The members of the Group 1 focused more on the Methods of the OOP because there are strong connections between the Knowledge of Methods and the Skills of Design (K.Methods - S.Design: 0.117) and Collaboration (K.Methods - S.Collaboration: 0.121), the Identity of Leader(K.Methods-I.Leader: 0.083) as well as the Epistemology of Design (K.Methods - E.Design: 0.084).There are also stronger connections between the Skill of Problem Solving and the Skill of Design(S.Design-S.ProblemSolving: 0.084), and the Skill of Collaboration(S.Collaboration-S.ProblemSolving: 0.088).That means that the members had the good knowledge and the skills needed to solve the exercise and collaborated over that.

The members of the Group 8 have stronger connections with the identity of Supportive and the Knowledge of Sorting. Specifically there were strong connections between Knowledge of Sorting and the Skill of Collaboration (K.Sorting - S.Collaboration: 0.141), the Identity of Supportive (K.Sorting- I.Supportive: 0.076) and the Skill of Design (K.Sorting-S.Design: 0.070). The connections from the identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.199), the Skill of Design (I.Supportive-S.Design: 0.0102) are also evident. Lastly there is a strong connection between the Epistemology of Data and the Skill of Collaboration (S.Collaboration-E.Data: 0.08). That means that the Group 1 has better skills in comparison to the other group but the members of the Group 8 focused more on the Sorting process in OOP .Lastly the identity of the Supportive appear prominent which means that the members of the Group 8 helped more each other in order to solve the exercise.

Comparison Group5 and Group8

In this case, the difference on the first and the second dimension is not significantly different::

- *Along the X axis, a two sample t test assuming unequal variance showed Group5 (mean=0.954, SD=0.761, N=4 was not statistically significantly different at the alpha=0.05 level from Group8 (mean=0.366, SD=0.368, N=3; t(4.505)= 1.351, p=0.241, Cohen's d=0.929).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group5 (mean=0.087, SD=0.918, N=4 was not statistically significantly different at the alpha=0.05 level from Group8 (mean=-0.538, SD=0.462, N=3; t(4.585)= 1.176, p=0.297, Cohen's d=0.812).*



Parametric

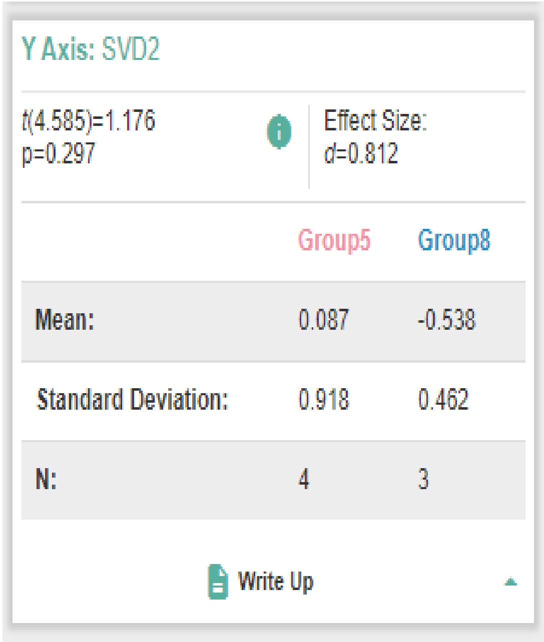
X Axis: SVD1

$t(4.505)=1.351$ p=0.241 | Effect Size: $d=0.929$

	Group5	Group8
Mean:	0.954	0.366
Standard Deviation:	0.761	0.368
N:	4	3

Write Up

Figure 3-97 The results of the Independent T-Test for the first dimension



Y Axis: SVD2

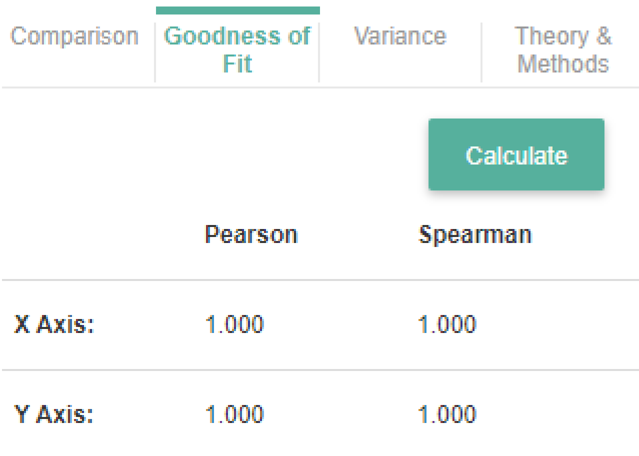
$t(4.585)=1.176$ p=0.297 | Effect Size: $d=0.812$

	Group5	Group8
Mean:	0.087	-0.538
Standard Deviation:	0.918	0.462
N:	4	3

Write Up

Figure 3-98 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-97 and Figure 4-98 show the independent t-test for the first and second dimension for Group5 and Group8. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are equal to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved.



Comparison | Goodness of Fit | Variance | Theory & Methods

Calculate

	Pearson	Spearman
X Axis:	1.000	1.000
Y Axis:	1.000	1.000

3-99 Pearson's and Spearman's R for the Goodness of Fit

The Figure 4-99 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two groups the two following networks for the two groups will appear. The Figure 4-100 shows the cognitive network for the Group5, the Figure 4-101 shows the Comparison Networks for Group5 and Group8 and the Figure 4-102 shows the cognitive network for the Group8.

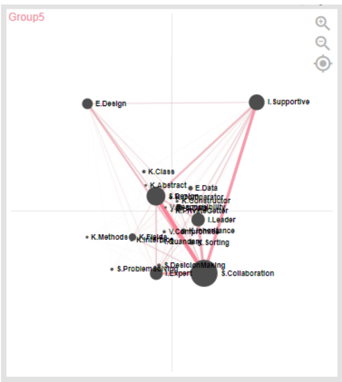


Figure 3-100 The Network of the centroid for Group5

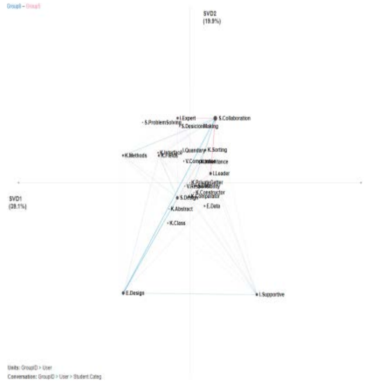


Figure 3-101 Comparison of Group5 and Group8

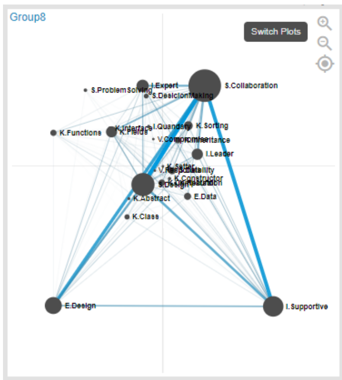


Figure 3-102 The Network of the centroid for Group8

There are no major differences between these two networks. The differences worth presenting are:

The members of the Group 1 have more prominent connections with the Identity nodes of Expert and Leader. More specifically there are strong connections in comparison to the other group between the identity of Expert and the Skills of Collaboration (I.Expert-S.Collaboration: 0.087) and Design (I.Expert-S.Design: 0.048) and the Identity of Leader (I.Expert-I.Leader: 0.056). There are also stronger connections between the Identity of Leader and the Skill of Design (I.Leader-S.Design: 0.064) and the Identity of Supportive (I.Leader-I.Supportive: 0.055).This group had more connections starting from the identities of Leader and Expert and one of the stronger connections was with the skill of design which means that there were members of this Group that have these two characteristics combined.

The members of the Group 8 have stronger connections with the Epistemology of Design and the Knowledge of Sorting. Specifically there were strong connections between Knowledge of Sorting and the Skill of Collaboration (K.Sorting - S.Collaboration: 0.065), the Identity of Supportive (K.Sorting- I.Supportive: 0.042). The connections between the Skill of Collaboration and the Epistemology of Design (E.Design-S.Collaboration: 0.101), the Knowledge of Comparator (K.Comparator-S.Collaboration: 0.054) and the Knowledge of Methods (K.Methods-S.Collaboration:

0.064) are also strong in comparison to the fifth group. Lastly there is a strong connection between the Epistemology of Design and the Skill of Design (E.Design-S.Design: 0.081) and the Identity Supportive (I.Supportive-E.Design: 0.058). That means that the Group 8 focused more on the Sorting process in OOP in comparison to the fifth group.

3.3.4.2 High/Low-to-High Category

The figure we see in ENA WebKit, after selecting the parameters, is the following (Figure 4-103), where we can see the centroids and the confidence intervals for ten student groups

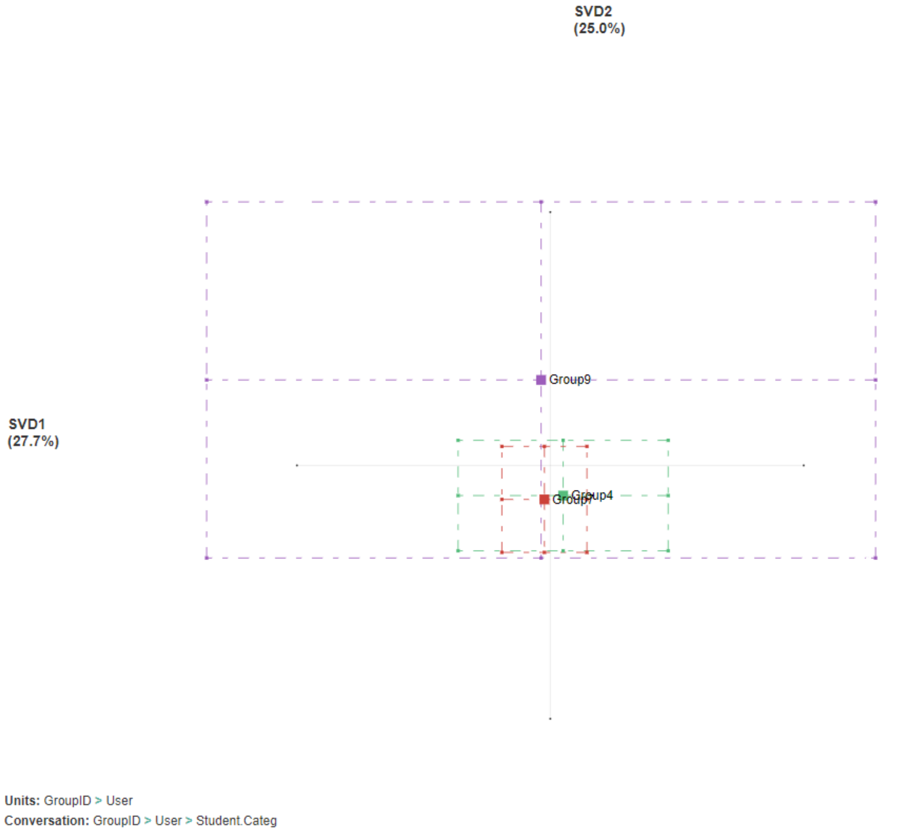


Figure 3-103 The Centroids of the Three Groups of the High/Low-to-High Category (GREEN: Group4, RED: Group7, PURPLE: Group9)

By default, the ENA space in which the centroids are displayed is determined by the first (x) and second (y) dimensions, the dimensions that represent the largest data variation. The numbers in parentheses next to the axis labels indicate the percentage of variation in the data relating to these dimensions. In this case, the dimension x (SVD1) represents 27.9% of the variance in the data and the dimension y (SVD2) represents 24.8%.

As shown in Figure 1, the three Groups are slightly different. To determine the difference more accurately, we can perform an independent samples t - test. To do this, simply select the two samples we want to compare from the drop-down menu on the left, in the tab "Stats". When we do this, we will see averages for the two samples, along with the t-score, the p value and Cohen's d, a measure of the magnitude of the effect.

Comparison Group4 and Group7

In this case, the difference on the first and the second dimension is not significantly different:

- *Along the X axis, a two sample t test assuming unequal variance showed Group4 (mean=-0.178, SD=0.917, N=4 was not statistically significantly different at the alpha=0.05 level from Group7 (mean=0.082, SD=0.371, N=4; t(3.957)= -0.526, p=0.627, Cohen's d=0.372).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group4 (mean=0.419, SD=0.482, N=4 was not statistically significantly different at the alpha=0.05 level from Group7 (mean=0.472, SD=0.463, N=4; t(5.990)= -0.160, p=0.878, Cohen's d=0.113).*

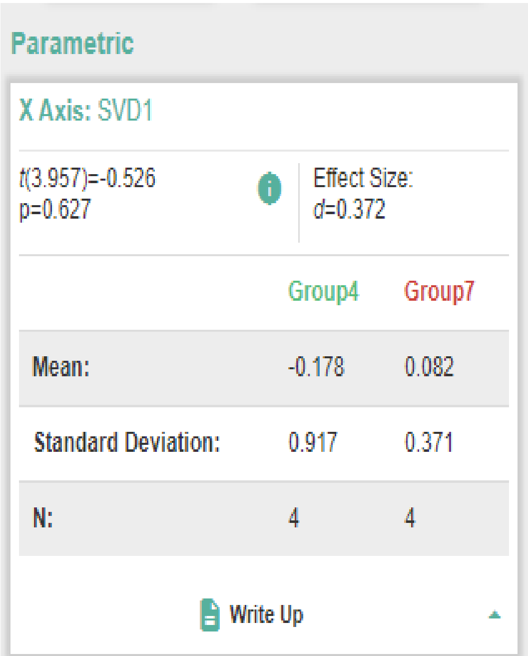


Figure 3-104 The results of the Independent T-Test for the first dimension

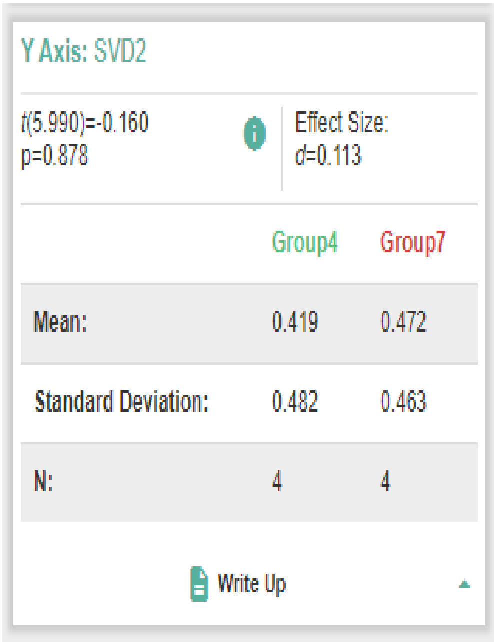


Figure 3-105 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-104 and Figure 4-105 show the independent t-test for the first and second dimension for Group4 and Group7. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson’s and Spearman’s r. In this case, both are close to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved.

Comparison | Goodness of Fit | Variance | Theory & Methods

Calculate

	Pearson	Spearman
X Axis:	1.000	1.000
Y Axis:	1.000	1.000

Figure 3-106 Pearson’s and Spearman’s R for the Goodness of Fit

The Figure 106 shows the goodness of fit for the data with the statistics of Pearson’s and Spearman’s R. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two groups the two following networks for the two groups will appear. The Figure 4-107 shows the cognitive network for the Group4, the Figure 4-108 shows the Comparison Networks for Group4 and Group7 and the Figure 4-109 shows the cognitive network for the Group7.

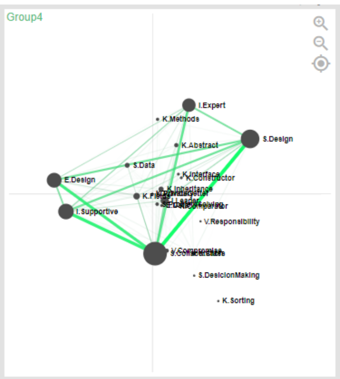


Figure 3-107 The Network of the centroid for Group4

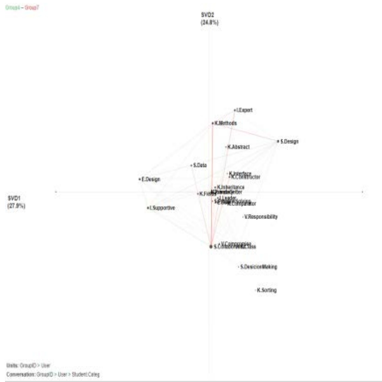


Figure 3-108 Comparison of Group4 and Group7

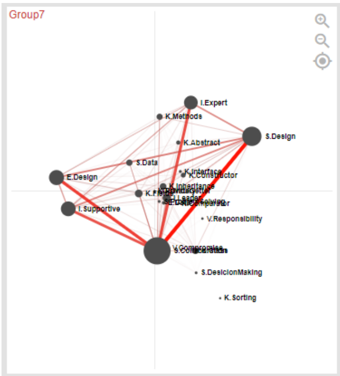


Figure 3-109 The Network of the centroid for Group7

There are no major differences between the two groups compared. The differences worth presenting are:

The members of the Group 4 have stronger connections with the Skill of Design and the identity of Supportive. Specifically there were strong connections between Skill of Design and the Epistemology of Design (E.Design- S.Design: 0.035) and the Identity of Expert (I.Expert- S.Design: 0.026). The connections from the identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.05), the skill of Design (I.Supportive- S.Design: 0.044) and the Epistemology of Design (I.Supportive-E.Design: 0.51). The members of the Group 4 have more prominent the identity of the Supportive which means, along with the strong connections with the epistemology of design, that the members of the Group 4 confirmed and helped more each other in order to solve the exercise.

The members of the Group 7 have stronger connections with the Skill of Collaboration and the Knowledge of Methods. Specifically there were strong connections between Skill of Collaboration and the Skill of Data (S.Collaboration-S.Data: 0.044), the Identity of Expert (S.Collaboration-I.Expert: 0.055) and the Knowledge of Constructor (S.Collaboration-K.Constructor: 0.059). The connections between the Knowledge of Methods and the Identity of Expert (I.Expert-K.Methods: 0.04), the Skill of Design (K.Methods-S.Design: 0.055) and the Skill of Collaboration (K.Methods-S.Collaboration: 0.089) are also strong. That means that in comparison the Group 4 has better design skills but the members of the Group 7 have stronger collaboration and knowledge of the Methods needed to solve the exercise.

Comparison Group4 and Group9

In this case, the difference on the first and the second dimension is not significantly different:

- *Along the X axis, a two sample t test assuming unequal variance showed Group4 (mean=-0.178, SD=0.917, N=4 was not statistically significantly different at the alpha=0.05 level from Group9 (mean=0.128, SD=1.870, N=3; t(2.727)= -0.261, p=0.813, Cohen's d=0.222).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group4 (mean=0.419, SD=0.482, N=4 was not statistically significantly different at the alpha=0.05 level from Group9 (mean=-1.188, SD=0.996, N=3; t(2.710)= 2.576, p=0.091, Cohen's d=2.194).*

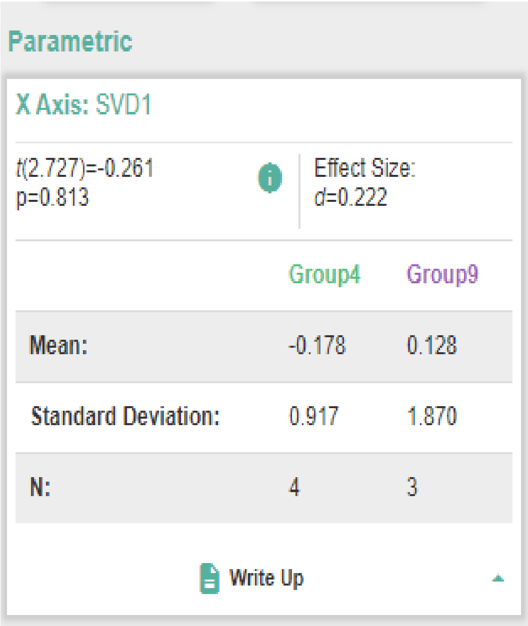


Figure 3-110 The results of the Independent T-Test for the first dimension

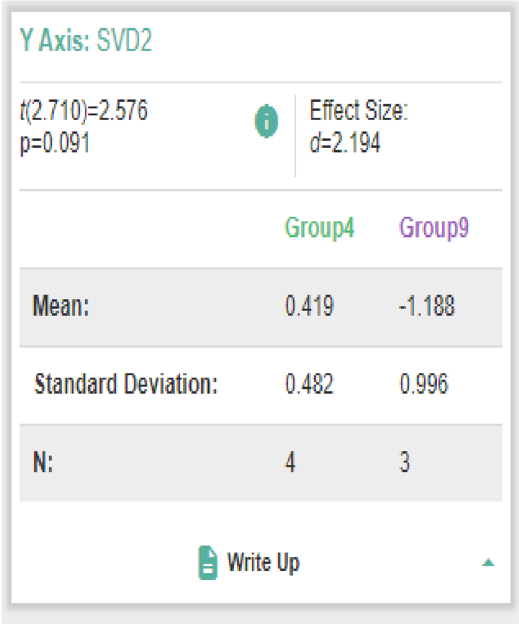
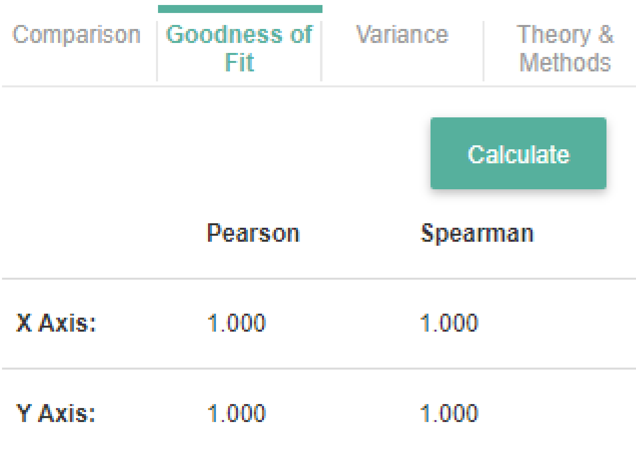


Figure 3-111 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-110 and Figure 4-111 show the independent t-test for the first and second dimension for Group4 and Group9. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are close to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved.



3-112 Pearson's and Spearman's R for the Goodness of Fit

The Figure 4-112 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these

two groups the two following networks for the two groups will appear. The Figure 4-113 shows the cognitive network for the Group4, the Figure 4-114 shows the Comparison Networks for Group4 and Group9 and the Figure 4-115 shows the cognitive network for the Group9.

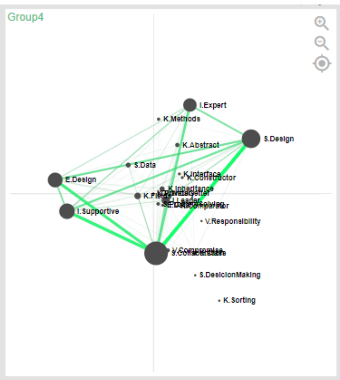


Figure 3-113 The Network of the centroid for Group4

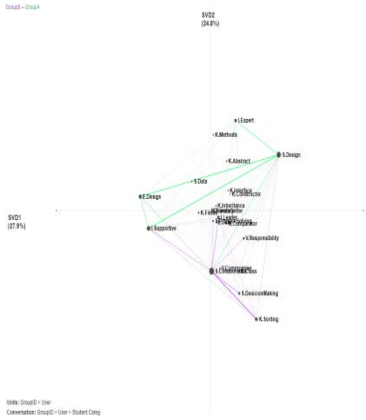


Figure 3-114 Comparison of Group4 and Group9

Figure 3-115 The Network of the centroid for Group9

There are no major differences between the two groups compared. The differences worth presenting are:

The members of the Group 4 have stronger connections with the Skill of Design and the identity of Supportive. Specifically there were strong connections between Skill of Design and the Epistemology of Design (E.Design- S.Design: 0.147), the Skill of Collaboration (S.Design- S.Collaboration: 0.105) and the Identity of Expert (I.Expert-S.Design: 0.12). The connections from the identity of Supportive and the Epistemology of Design (I.Supportive-E.Design: 0.109), the skill of Design (I.Supportive- S.Design: 0.152) and the Epistemology of Design (I.Supportive-E.Design: 0.51). The members of the Group 4 have more prominent the identity of the Supportive which means, along with the strong connections with the epistemology of design, that the members of the Group 4 confirmed and helped more each other in order to solve the exercise.

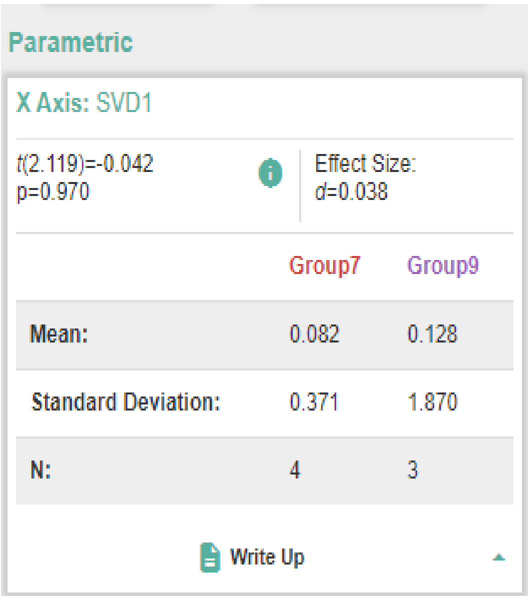
The members of the Group 9 have stronger connections with the Skill of Collaboration. Specifically there were strong connections between Skill of Collaboration and the Knowledge of Class in Java (K.Class-S.Collaboration: 0.1) and the Knowledge of Sorting (K.Sorting-S.Collaboration: 0.147). The connections from the Skill of Collaboration and the Identity of Leader (I.Leader-S.Collaboration:0.107) and the Skill of Problem Solving (S.ProblemSolving-S.Collaboration: 0.072) are also evident. Also there are strong connections between the Skill of Collaboration and the Value of Responsibility (S.Collaboration-V.Responsibility: 0.08) and Skill of Decision Making

(S.Collaboration-S.DesicionMaking: 0.107). Lastly in comparison to the previous group there are strong connections between the Identity of Supportive and the Knowledge of Sorting (K.Sorting- I.Supportive: 0.06) and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.087). That means that the members of the Group 9 have prominent the skill of collaboration thus they collaborated better with each other to solve the exercise.

Comparison Group7 and Group9

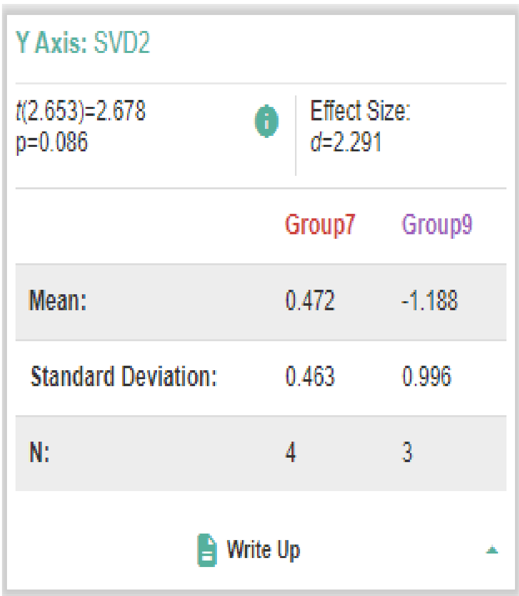
In this case, the difference on the first and the second dimension is not significantly different:

- *Along the X axis, a two sample t test assuming unequal variance showed Group7 (mean=0.082, SD=0.371, N=4 was not statistically significantly different at the alpha=0.05 level from Group9 (mean=0.128, SD=1.870, N=3; t(2.119)= -0.042, p=0.970, Cohen's d=0.038).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group7 (mean=0.472, SD=0.463, N=4 was not statistically significantly different at the alpha=0.05 level from Group9 (mean=-1.188, SD=0.996, N=3; t(2.653)= 2.678, p=0.086, Cohen's d=2.291).*



	Group7	Group9
Mean:	0.082	0.128
Standard Deviation:	0.371	1.870
N:	4	3

Figure 3-116 The results of the Independent T-Test for the first dimension



	Group7	Group9
Mean:	0.472	-1.188
Standard Deviation:	0.463	0.996
N:	4	3

Figure 3-117 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-116 and Figure 4-117 show the independent t-test for the first and second dimension for Group7 and Group9. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson’s and Spearman’s r. In this case, both are close to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved.

Comparison | **Goodness of Fit** | Variance | Theory & Methods

Calculate

	Pearson	Spearman
X Axis:	1.000	1.000
Y Axis:	1.000	1.000

Figure 3-118 Pearson's and Spearman's R for the Goodness of Fit

The Figure 4-118 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two groups the two following networks for the two groups will appear. The Figure 4-119 shows the cognitive network for the Group7, the Figure 4-120 shows the Comparison Networks for Group7 and Group9 and the Figure 4-121 shows the cognitive network for the Group9.

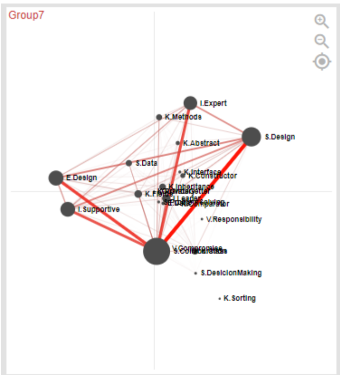


Figure 3-119 The Network of the centroid for Group7

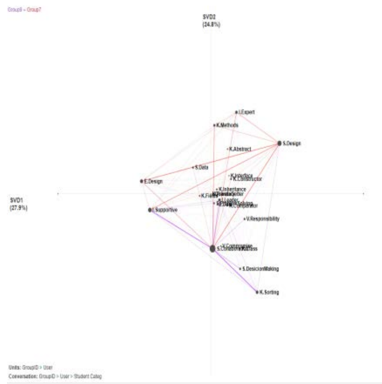


Figure 3-120 Comparison of Group7 and Group9

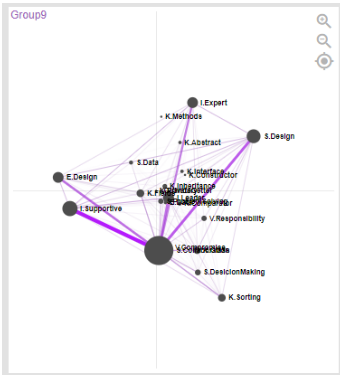


Figure 3-121 The Network of the centroid for Group9

There are no major differences between the two groups. The differences worth presenting are:

The members of the Group 7 have stronger connections with the Skill of Collaboration and the Skill of Design. Specifically there were strong connections between Skill of Collaboration and the Skill of Design (S.Design- S.Collaboration:

0.124), the Identity of Expert (S.Collaboration-I.Expert: 0.088), the Knowledge of Methods (K.Methods-S.Collaboration: 0.11) and the Epistemology of Design (S.Collaboration-E.Design: 0.085). The connections between the Skill of Design and the Identity of Supportive (I.Supportive- S.Design: 0.108), the Epistemology of Design (S.Design- E.Design: 0.112), the Knowledge of Methods (K.Methods-S.Design: 0.072) and Identity of Expert (S.Design-I.Expert: 0.094) are also prominent. That interpretation can be the higher skills of collaboration and design from the members of the Group7.

The members of the Group 9 have stronger connections with the Skill of Collaboration in comparison to the previous group. Specifically there were strong connections between the Skill of Collaboration and the Identity of Leader (I.Leader-S.Collaboration:0.109) and the Identity of Supportive (I.Supportive- S.Collaboration: 0.137). Also there are strong connections between the Skill of Collaboration and the Skill of Problem Solving (S.ProblemSolving-S.Collaboration: 0.083) and Skill of Decision Making (S.Collaboration-S.DesicionMaking: 0.087). Lastly there are strong connections between the Knowledge of Sorting and the Skill of Collaboration (K.Sorting-S.Collaboration: 0.147) and the Skill of Design (K.Sorting-S.Design: 0.064). That means that the members of the Group 9 have prominent the skill of collaboration thus they collaborated better with each other to solve the exercise. That means that in comparison the Group 7 has better design skills but the members of the Group 9 have stronger collaboration and knowledge of the sorting techniques in OOP, needed to solve the exercise.

3.3.4.3 High/Low-to-Low Category

The figure we see in ENA WebKit, after selecting the parameters, is the following (Figure 4-122), where we can see the centroids and the confidence intervals for three student groups

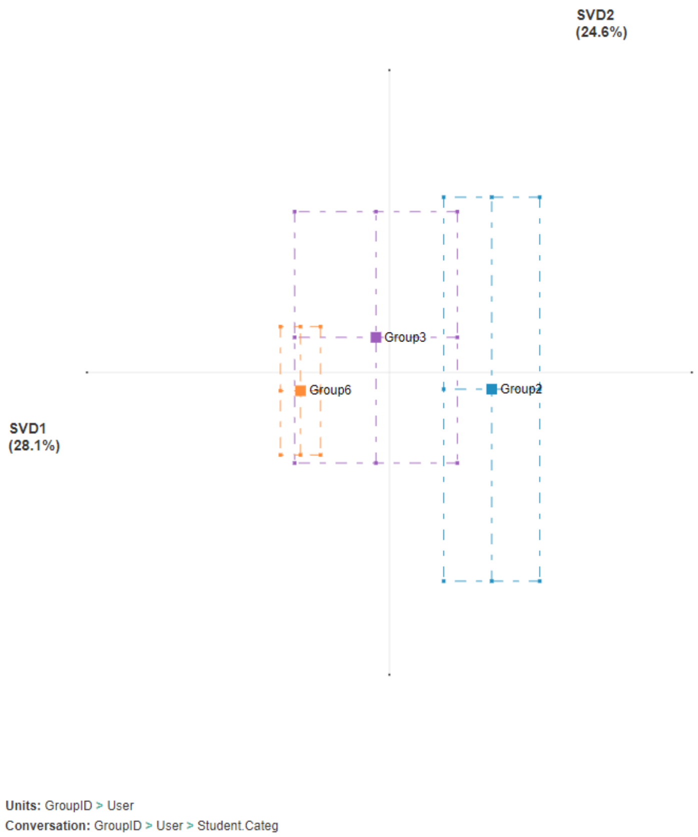


Figure 3-122 The Centroids of the Three Groups of High/Low-to-Low Category

By default, the ENA space in which the centroids are displayed is determined by the first (x) and second (y) dimensions, the dimensions that represent the largest data variation. The numbers in parentheses next to the axis labels indicate the percentage of variation in the data relating to these dimensions. In this case, the dimension x (SVD1) represents 28.1% of the variance in the data and the dimension y (SVD2) represents 24.6%.

As shown in Figure 122, the three Groups are slightly different. To determine the difference more accurately, we can perform an independent samples t - test. To do this, simply select the two samples we want to compare from the drop-down menu on the left, in the tab "Stats". When we do this, we will see averages for the two samples, along with the t-score, the p value and Cohen's d, a measure of the magnitude of the effect.

Comparison Group2 and Group3

In this case there is not a significant difference on the second dimension but there is a significant difference on the first one:

- *Along the X axis, a two sample t test assuming unequal variance showed Group2 (mean=1.655, SD=0.488, N=4 was statistically significantly different* *at the alpha=0.05 level from Group3 (mean=-0.217, SD=0.827, N=4; t(4.864)= -3.898, p=0.012, Cohen's d=2.756).*

- *Along the Y axis, a two sample t test assuming unequal variance showed Group2 (mean=-0.271, SD=1.953, N=4 was not statistically significantly different at the alpha=0.05 level from Group3 (mean=0.568, SD=1.279, N=4; t(5.173)= -0.718, p=0.504, Cohen's d=0.508).*

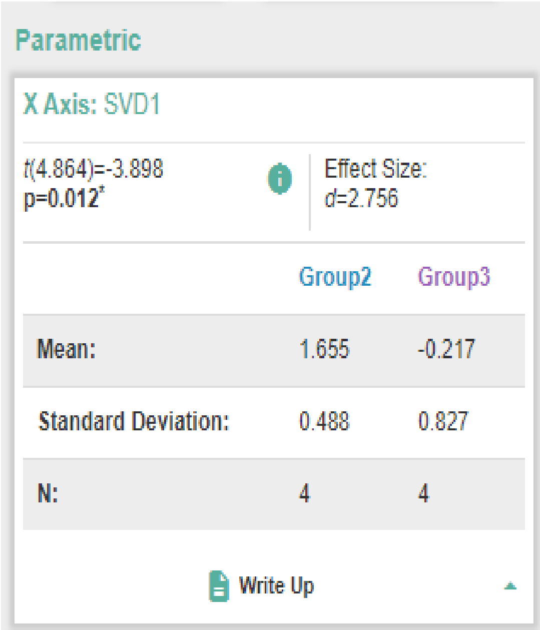


Figure 3-123 The results of the Independent T-Test for the first dimension

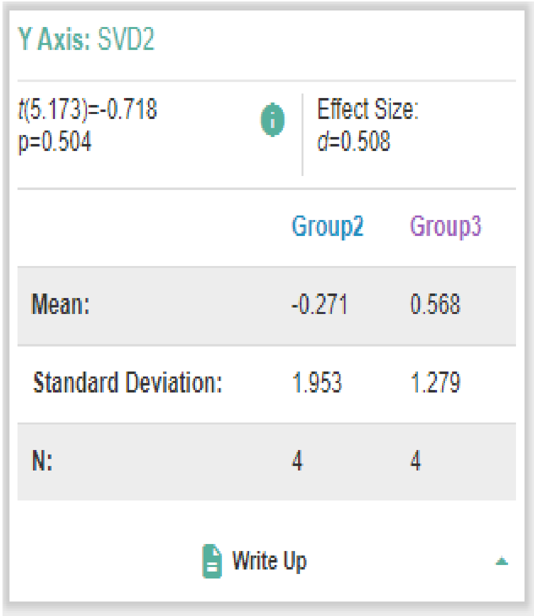


Figure 3-124 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-123 and Figure 4-124 show the independent t-test for the first and second dimension for Group2 and Group3. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are close to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved.

Comparison | Goodness of Fit | Variance | Theory & Methods

Calculate

	Pearson	Spearman
X Axis:	1.000	1.000
Y Axis:	1.000	1.000

3-125 Pearson’s and Spearman’s R for the Goodness of Fit

The Figure 4-125 shows the goodness of fit for the data with the statistics of Pearson’s and Spearman’s R. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two groups the two following networks for the two groups will appear. The Figure 4-126 shows the cognitive network for the Group2, the Figure 4-127 shows the Comparison Networks for Group2 and Group3 and the Figure 4-128 shows the cognitive network for the Group3.

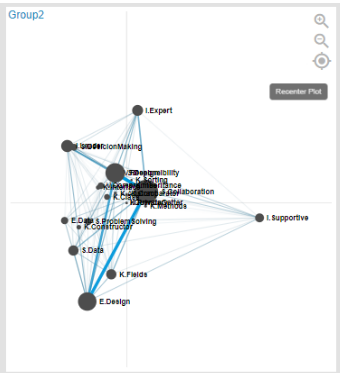


Figure 3-126 The Network of the centroid for Group2

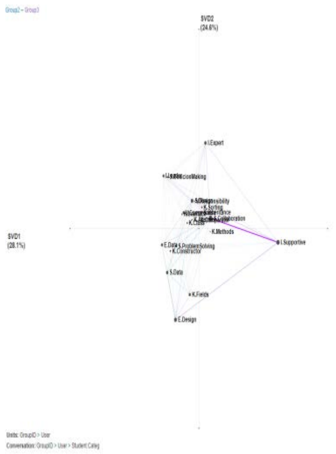


Figure 3-127 Comparison of Group2 and Group3

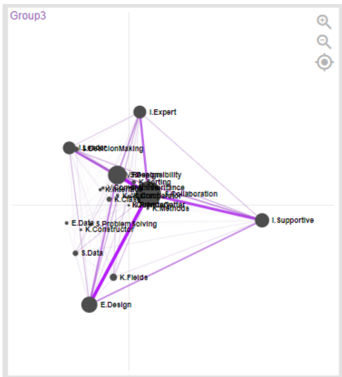


Figure 3-128 The Network of the centroid for Group3

There are no major differences between the two Groups. The differences worth presenting are:

The members of the Group 2 have stronger connections with the Epistemology category codes which mean that they tend to confirm more often the answers of the other members. Specifically there were strong connections between Epistemology of Design and the Skill of Design (S.Design-E.Design: 0.065) and the Epistemology of Data

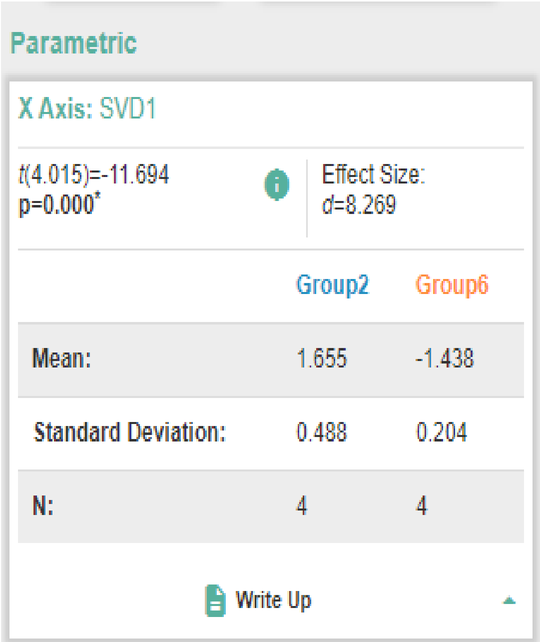
(E.Data - E.Design: 0.07) but there was also strong connection between the skill of Collaboration and the skill of Data (S.Collaboration - S.Data: 0.068). Lastly there are strong connections between the Epistemology of Data and the Knowledge of Fields (K.Fields-E.Data: 0.042) and the Skill of Collaboration (S.Collaboration-E.Data: 0.057).

The members of the Group 3 have stronger connections with the Identity of Supportive. Specifically there were strong connections between Identity of Supportive and the Epistemology of Design (I.Supportive-E.Design: 0.105), the Identity of Expert (I.Supportive-I.Expert: 0.064) and the Skills of Collaboration (I.Supportive-S.Collaboration: 0.232) and Design (I.Supportive-S.Design: 0.088). Lastly another strong connection in comparison to the other group is the one between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.078). The results of the comparison show that the identity of Supportive is the most important in the Group3 and that the members of the Group2 tend to confirm each other more.

Comparison Group2 and Group6

In this case there is not a significant difference on the second dimension but there is a significant difference on the first one:

- *Along the X axis, a two sample t test assuming unequal variance showed Group2 (mean=1.655, SD=0.488, N=4 was statistically significantly different at the alpha=0.05 level from Group6 (mean=-1.438, SD=0.204, N=4; t(4.015)= -11.694, p=0.000, Cohen's d=8.269).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group2 (mean=-0.271, SD=1.953, N=4 was not statistically significantly different at the alpha=0.05 level from Group6 (mean=-0.297, SD=0.653, N=4; t(3.662)= 0.026, p=0.981, Cohen's d=0.018).*



Parametric

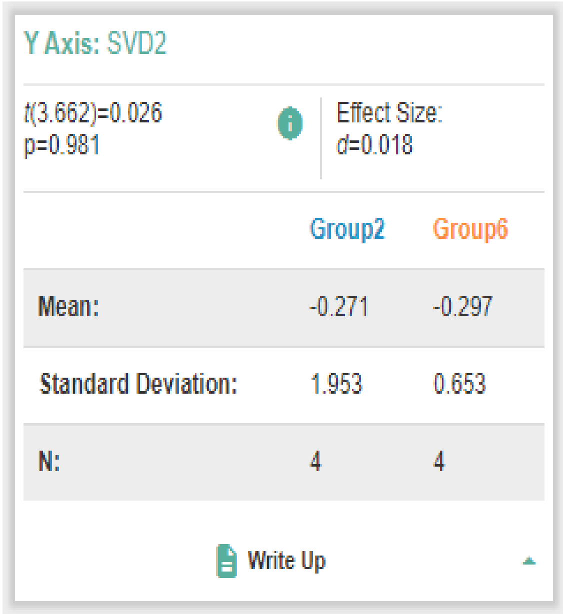
X Axis: SVD1

t(4.015)=-11.694 p=0.000* | Effect Size: *d*=8.269

	Group2	Group6
Mean:	1.655	-1.438
Standard Deviation:	0.488	0.204
N:	4	4

Write Up

Figure 3-129 The results of the Independent T-Test for the first dimension



Y Axis: SVD2

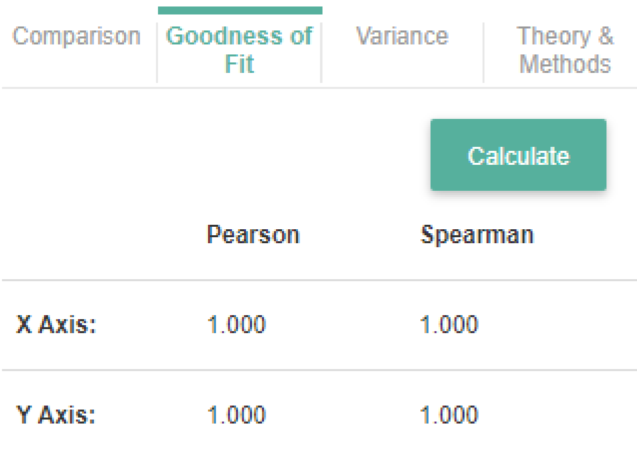
t(3.662)=0.026 p=0.981 | Effect Size: *d*=0.018

	Group2	Group6
Mean:	-0.271	-0.297
Standard Deviation:	1.953	0.653
N:	4	4

Write Up

Figure 3-130 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-129 and Figure 4-130 show the independent t-test for the first and second dimension for Group2 and Group6. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are close to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved.



Comparison | Goodness of Fit | Variance | Theory & Methods

Calculate

	Pearson	Spearman
X Axis:	1.000	1.000
Y Axis:	1.000	1.000

Figure 3-131 Pearson's and Spearman's R for the Goodness of Fit

The Figure 4-131 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these

two groups the two following networks for the two groups will appear. The Figure 4-132 shows the cognitive network for the Group2, the Figure 4-133 shows the Comparison Networks for Group2 and Group6 and the Figure 4-134 shows the cognitive network for the Group6.

Figure 3-132 The Network of the centroid for Group2

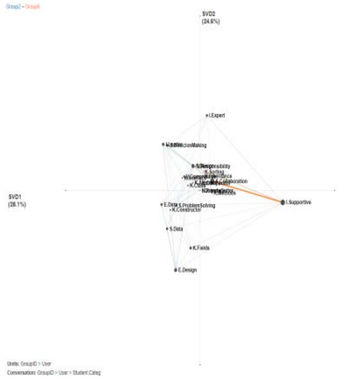


Figure 3-133 Comparison of Group2 and Group6

Figure 3-134 The Network of the centroid for Group6

There are many differences between the two Groups. The differences worth presenting are:

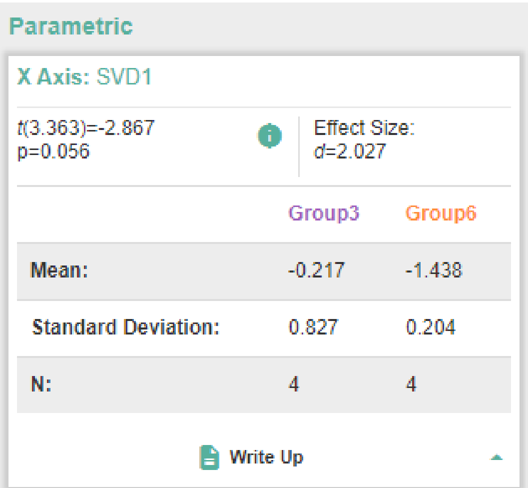
The members of the Group2 have stronger connections with the Epistemology and the Skill category codes. Specifically there were strong connections between Epistemology of Design and the Skill of Design (S.Design-E.Design: 0.087) and the Epistemology of Data (E.Data - E.Design: 0.079). There are also strong connections between the Skill of Collaboration and the Identity of Leader (I.Leader-S.Collaboration: 0.077) and the Skills of Decision Making (S.DesicionMaking-S.Collaboration: 0.081) and Problem Solving (S.ProblemSolving-S.Collaboration: 0.083).

The members of the Group6 have stronger connections with the Identity of Supportive. Specifically there were strong connections between Identity of Supportive and the Epistemology of Design (I.Supportive-E.Design: 0.073), the Identity of Expert (I.Supportive-I.Expert: 0.099) and the Skills of Collaboration (I.Supportive-S.Collaboration: 0.384) and Design (I.Supportive-S.Design: 0.168). Lastly another strong connection in comparison to the other group is the one between the Identity of Supportive and the Knowledge of Fields (I.Supportive-K.Fields: 0.079). The results of the comparison show that the identity of Supportive is the most important in the Group6 and that the members of the Group2 tend to confirm each other more.

Comparison Group3 and Group6

In this case, the difference on the first and the second dimension is not significantly different::

- *Along the X axis, a two sample t test assuming unequal variance showed Group3 (mean=-0.217, SD=0.827, N=4 was not statistically significantly different at the alpha=0.05 level from Group6 (mean=-1.438, SD=0.204, N=4; t(3.363)= -2.867, p=0.056, Cohen's d=2.027).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Group3 (mean=0.568, SD=1.279, N=4 was not statistically significantly different at the alpha=0.05 level from Group6 (mean=-0.297, SD=0.653, N=4; t(4.465)= 1.205, p=0.288, Cohen's d=0.852).*



Parametric

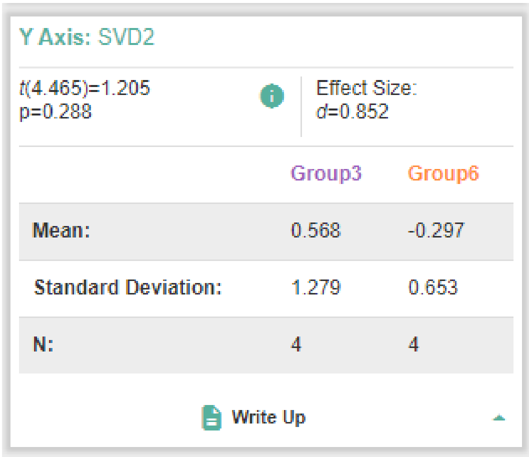
X Axis: SVD1

t(3.363)=-2.867 p=0.056 | Effect Size: d=2.027

	Group3	Group6
Mean:	-0.217	-1.438
Standard Deviation:	0.827	0.204
N:	4	4

Write Up

Figure 3-135 The results of the Independent T-Test for the first dimension



Y Axis: SVD2

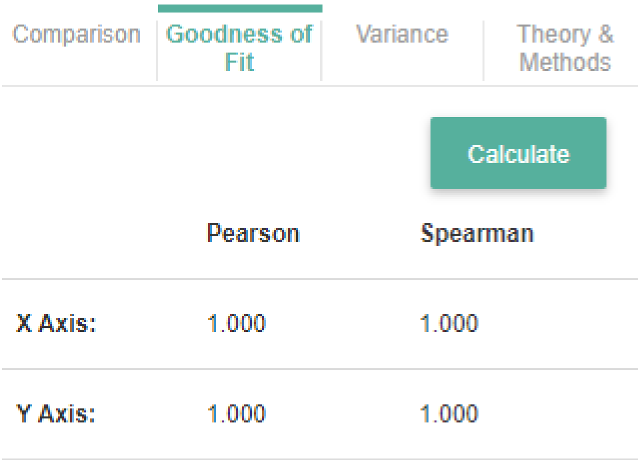
t(4.465)=1.205 p=0.288 | Effect Size: d=0.852

	Group3	Group6
Mean:	0.568	-0.297
Standard Deviation:	1.279	0.653
N:	4	4

Write Up

Figure 3-136 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-135 and Figure 4-136 show the independent t-test for the first and second dimension for Group3 and Group6. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are close to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved.



Comparison | Goodness of Fit | Variance | Theory & Methods

Calculate

	Pearson	Spearman
X Axis:	1.000	1.000
Y Axis:	1.000	1.000

Figure 3-137 Pearson's and Spearman's R for the Goodness of Fit

The Figure 4-137 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two groups the two following networks for the two groups will appear. The Figure 4-138 shows the cognitive network for the Group3, the Figure 4-139 shows the Comparison Networks for Group3 and Group6 and the Figure 4-140 shows the cognitive network for the Group6.

Figure 3-138 The Network of the centroid for Group3

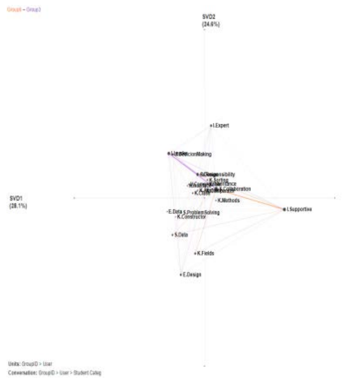


Figure 3-139 Comparison of Group3 and Group6

Figure 3-140 The Network of the centroid for Group6

There are no major differences between the two groups. The differences worth presenting are:

The members of the Group 3 have stronger connections with the Identity of Leader. Specifically there were strong connections between Identity of Leader and the Epistemology of Design (I.Leader-E.Design: 0.053), the Identity of Expert (I.Supportive-I.Leader: 0.064) and the Skill of Design (I.Leader-S.Design: 0.092). Lastly strong connections in comparison to the other group are the ones between the Skill of Collaboration and the Identity of Leader (I.Leader-S.Collaboration: 0.122) and the Skill of Decision Making (S.Collaboration-S.DesicionMaking: 0.064).

The members of the Group6 have stronger connections with the Identity of Supportive and the Skill of Collaboration. Specifically there were strong connections between Identity of Supportive and the Epistemology of Design (I.Supportive-E.Design: 0.080), the Skill of Data (I.Supportive-S.Data: 0.066) and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.152). There are also strong connections between the Collaboration and Skill of Data (S.Collaboration-S.Data: 0.068) and the Knowledge of Fields (s.Collaboration-K.Fields: 0.096). The results of the comparison show that the identity of Supportive is the most important in the Group6 and that the members of the Group2 tend to confirm each other more.

3.3.5 RQ4. Is there a significant difference between the discourse networks of the categories of the Groups based on the fundamental OOP concepts they used?

In Case 2 we examine the knowledge and the usage of fundamentals concepts of Object Oriented Programming such as Abstract, Inheritance and Comparator from the students to solve an Object Oriented Exercise using Java based on their grade on the course Object Oriented Programming.

The research question that is going to be answered in this subsection is:

RQ4. Is there a significant difference between the discourse networks of the categories of the Groups based on the fundamental OOP concepts they used?

The units, the codes and the conversation that were chosen are the following:

- Units: Inh.Abs.Com , GroupID, UserID
- Conversation: UserID,GroupID ,Student.Categ, Inh.Abs.Com
- Stanza Window: Moving Window Consisted of 4 Lines
- Codes: All the codes given in the file
- Comparison: Inh.Abs.Com

The figure we see in ENA WebKit, after selecting the parameters, is the following (Figure 4-141), where we can see the centroids and the confidence intervals for the four student categories:

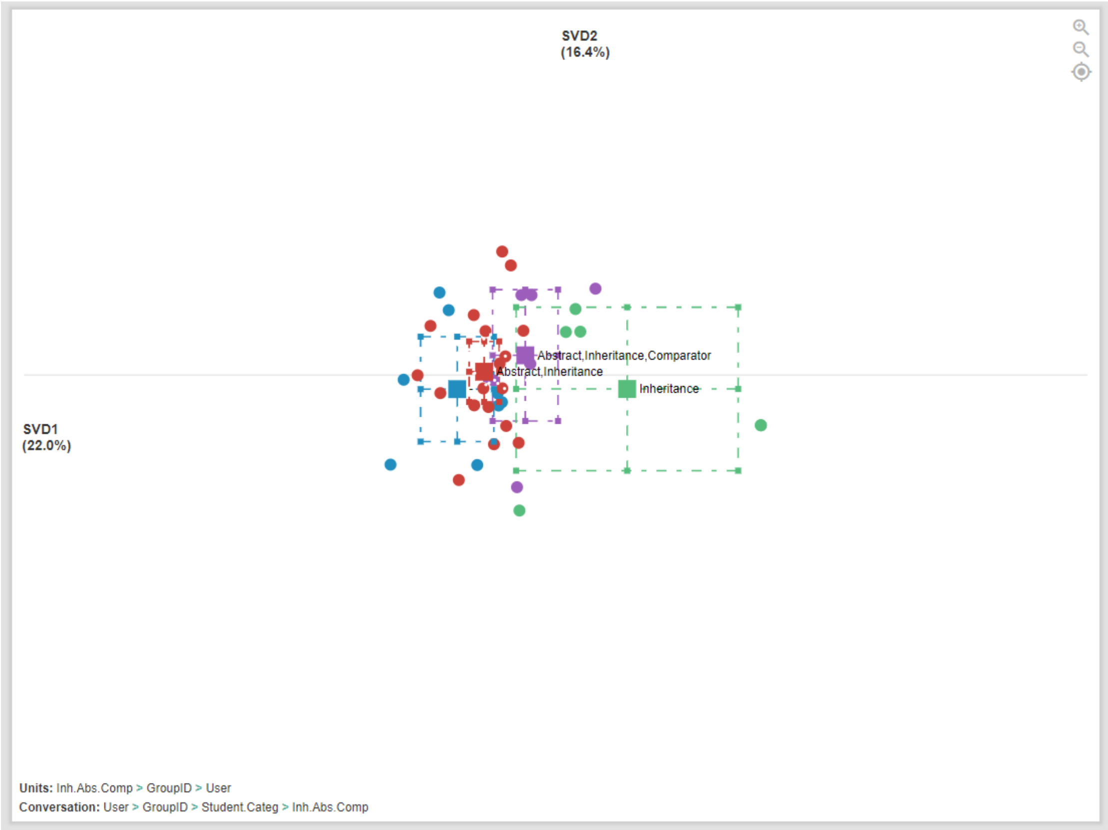


Figure 3-141 The Centroids of the Four Group Categories (RED: Abstract, Inheritance, GREEN: Inheritance, PURPLE: Abstract, Inheritance, Comparator, BLUE: None)

By default, the ENA space in which the centroids are displayed is determined by the first (x) and second (y) dimensions, the dimensions that represent the largest data variation. The numbers in parentheses next to the axis labels indicate the percentage of variation in the data relating to these dimensions. In this case, the dimension x (SVD1) represents 22.0% of the variance in the data and the dimension y (SVD2) represents 16.4%.

As shown in Fig. 141, the four Categories networks are slightly different. To determine the difference more accurately, we can perform an independent samples t - test. To do this, simply select the two samples we want to compare from the drop-down menu on the left, in the tab "Stats". When we do this, we will see averages for the two samples, along with the t-score, the p value and Cohen's d, a measure of the magnitude of the effect. In order to analyze them further we will present the differences of the connections between the two groups. This difference is calculated by subtracting the weight of the edge of one network from the other. The numbers in the brackets is the numeric differences of each edge between the two Groups or Categories.

3.3.5.1 Comparison of the Groups that used the fundamentals Concepts of Abstract and Inheritance and the Groups that used the fundamentals Concepts of Abstract, Inheritance and Comparator

In this case there is not a significant difference on the second dimension but there is a significant difference on the first one:

- *Along the X axis, a two sample t test assuming unequal variance showed Abstract,Inheritance (mean=0.548, SD=0.693, N=18 was statistically significantly different at the alpha=0.05 level from Abstract,Inheritance,Comparator (mean=-0.397, SD=0.812, N=7; t(9.611)= -2.718, p=0.022, Cohen's d=1.302).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Abstract,Inheritance (mean=0.076, SD=1.397, N=18 was not statistically significantly different at the alpha=0.05 level from Abstract,Inheritance,Comparator (mean=0.451, SD=1.628, N=7; t(9.646)= -0.537, p=0.603, Cohen's d=0.257).*

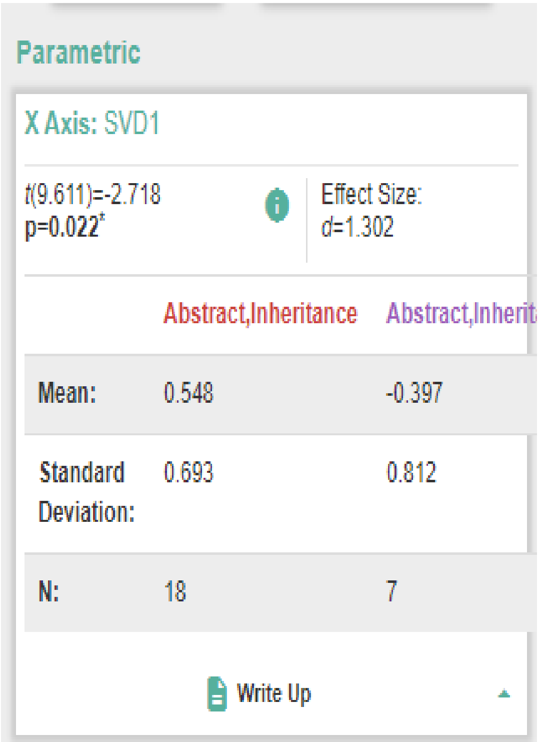


Figure 3-142 The results of the Independent T-Test for the first dimension

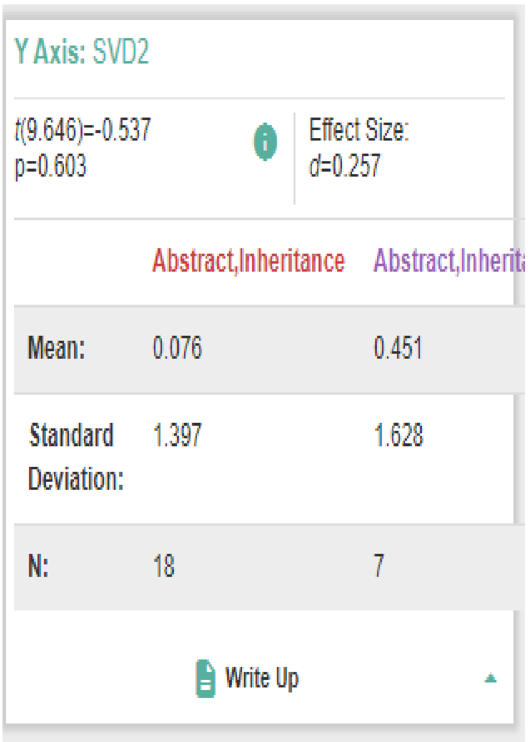


Figure 3-143 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-142 and Figure 4-143 show the independent t-test for the first and second dimension for the categories Abstract, Inheritance and Abstract, Inheritance, Comparator. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are equal to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved.

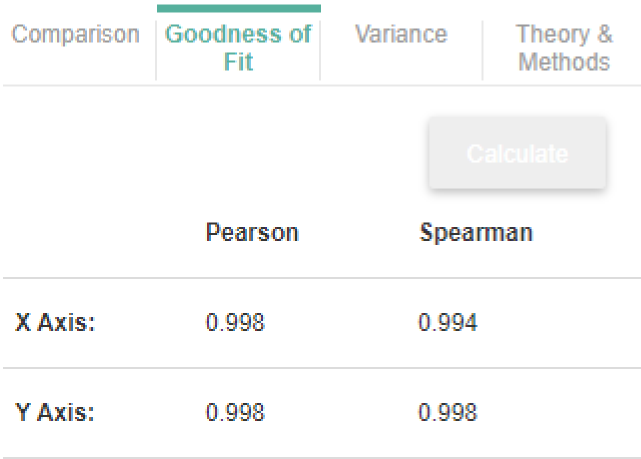


Figure 3-144 Pearson's and Spearman's R for the Goodness of Fit

The Figure 4-144 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two categories the two following networks for the two categories will appear. The Figure 4-145 shows the cognitive network for the Abstract, Inheritance, Comparator category and the Figure 4-146 shows the cognitive network for the Abstract, Inheritance category.

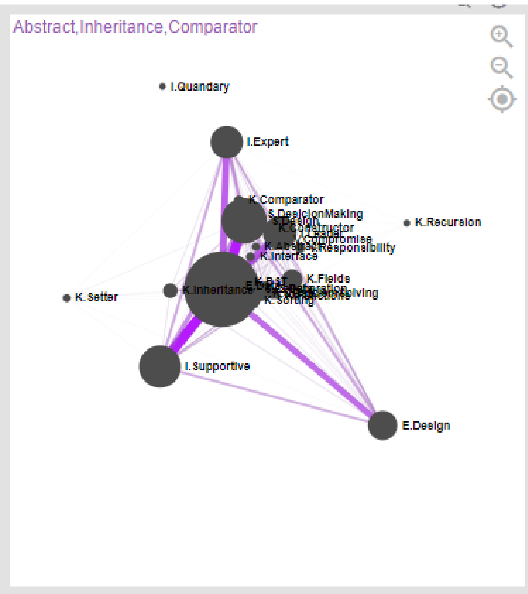


Figure 3-145 The Network of the centroid of groups used Abstract, Inheritance, Comparator

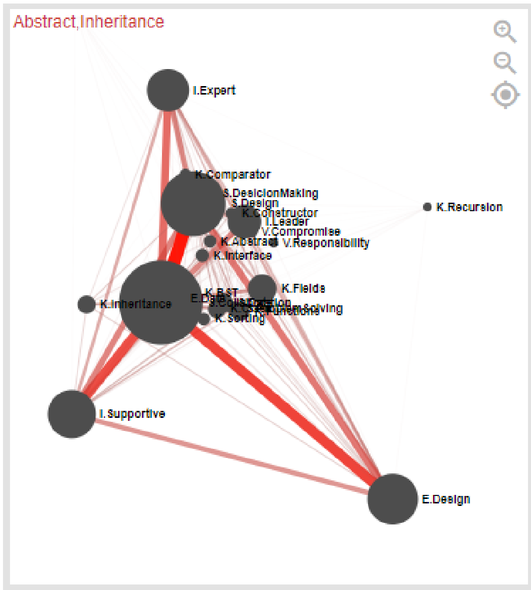


Figure 3-146 The Network of the centroid of groups used Abstract, Inheritance

There are many differences between the two networks analyzed. The differences worth presenting are:

The members of the groups that used the concepts of Abstract, Inheritance and Comparator have more connections between the Skill and Knowledge nodes. More specifically the stronger connections are formed between the Epistemology of Design and the Skill of Design (S.Design- E.Design: 0.114), the Skill of Collaboration (S.Collaboration-E.Design: 0.093) and the Identity of Supportive (I.Supportive-E.Design: 0.064). There are also strong connections between the Knowledge of Functions and the Skill of Collaboration (K.Functions-S.Collaboration: 0.069) and the Skill of Design (K.Functions-S.Design: 0.051). That means that the members had the good skills and knowledge of the methods needed and they tend to confirm the proposition of one another more often.

The members of the groups that used the concepts of Abstract, Inheritance and Comparator have more connections starting from the Identity of Supportive and the Skill of Collaboration. In particular there are strong connections between the Skill of Collaboration and the Identity of Supportive (I.Supportive-S.Collaboration: 0.083), the Identity of Leader (I.Leader-S.Collaboration: 0.105) and the Knowledge of Sorting (K.Sorting-S.Collaboration: 0.074).There are also strong connections between the Identities of Supportive and Leader (I.Supportive-I.Leader: 0.055) and between the Skills of Sorting and Design (K.Sorting-S.Design(0.034). That means this category has distinctive roles, better collaboration and knowledge of Sorting. Also the identity of Supportive appears prominent which means that they helped more each other in order to solve the exercise.

3.3.5.2 Comparison of the Groups that used the fundamentals Concepts of Abstract and Inheritance and the Groups that used the fundamentals Concept of Inheritance

In this case there is not a significant difference on the second dimension but there is a significant difference on the first one:

- *Along the X axis, a two sample t test assuming unequal variance showed Abstract,Inheritance (mean=0.548, SD=0.693, N=18 was statistically significantly different at the alpha=0.05 level from Inheritance (mean=-2.735, SD=2.430, N=6; t(5.273)= -3.266, p=0.021, Cohen's d=2.509).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Abstract,Inheritance (mean=0.076, SD=1.397, N=18 was not statistically significantly different at the alpha=0.05 level from Inheritance (mean=-0.321, SD=1.786, N=6; t(7.159)= 0.496, p=0.635, Cohen's d=0.265).*

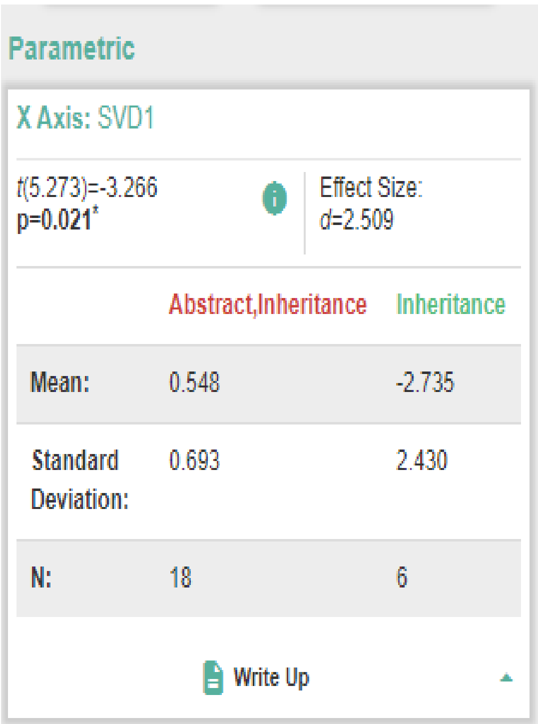


Figure 3-147 The results of the Independent T-Test for the first dimension

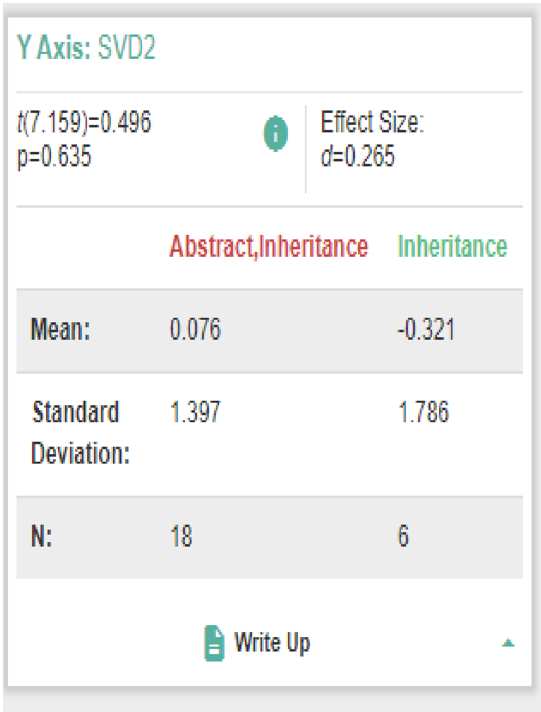


Figure 3-148 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-147 and Figure 4-148 show the independent t-test for the categories Abstract, Inheritance and Inheritance. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are equal to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved.

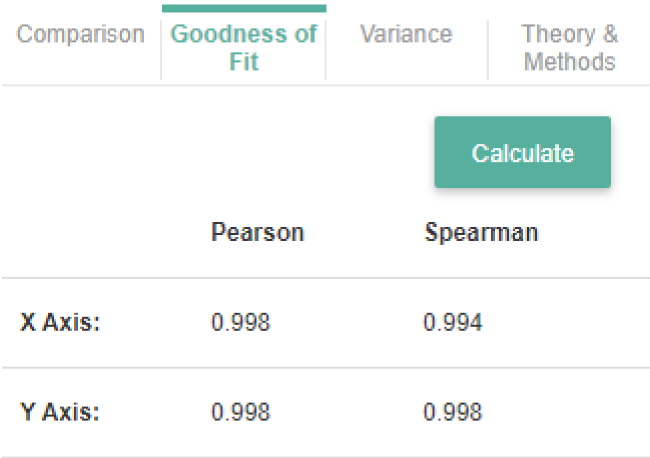


Figure 3-149 Pearson's and Spearman's R for the Goodness of Fit

The Figure 4-149 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two categories the two following networks for the two categories will appear. The Figure 4-150 shows the cognitive network for the Inheritance category and the Figure 4-151 shows the cognitive network for the Abstract, Inheritance category.

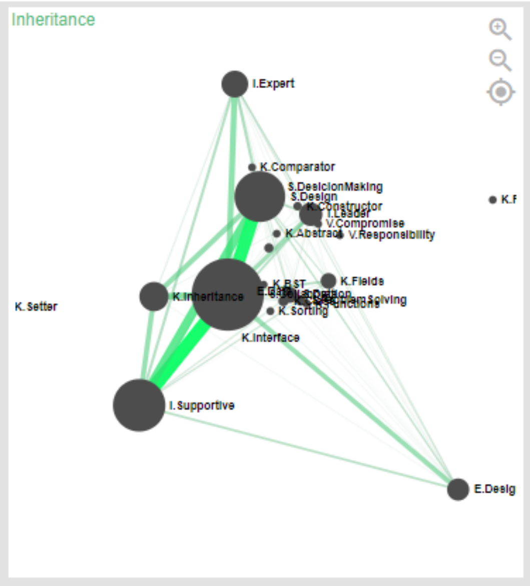


Figure 3-150 The Network of the centroid for Inheritance

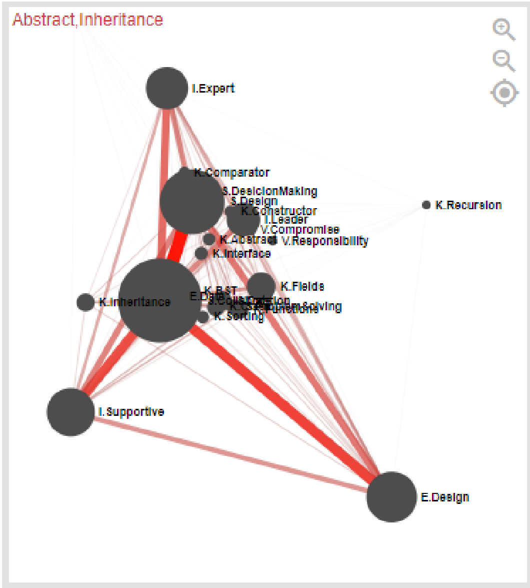


Figure 3-151 The Network of the centroid for Abstract, Inheritance

There are many differences between the two groups analyzed. The differences worth presenting are:

The members of the groups that used the concepts of Abstract and Inheritance have more connections with the Skill and Epistemology nodes. More specifically the stronger connections are formed between the Epistemology of Design and the Skill of Design (S.Design- E.Design: 0.174) and the Identity of Supportive (I.Supportive-E.Design: 0.086). There are also strong connections between the Skill of Collaboration and the Knowledge of Functions (S.Collaboration-K.Functions: 0.073) and Epistemology of Design (S.Collaboration-E.Design: 0.183). Also there are strong connections between the Identity of Expert and the Skill of Design (I.Expert-S.Design: 0.089) and the Epistemology of Design (I.Expert-E.Design: 0.069). That means that the members had the design skills needed and they tend to confirm the proposition of one another more often.

The members of the groups that used the concept of Inheritance have more connections starting from the Identity of Supportive and the Knowledge of Inheritance. In particular there are strong connections between the Identity of Supportive and the Skill of Design (I.Supportive-S.Design: 0.095), the Skill of Collaboration (I.Supportive-S.Collaboration: 0.132) and the Knowledge of Inheritance (I.Supportive- K.Inheritance: 0.139). There are also strong connections between the Knowledge of Inheritance and the Skill of Design (K.Inheritance-S.Design: 0.137) and the Skill of Collaboration (K.Inheritance-S.Collaboration: 0.156). That means this category focused more on

Inheritance which can be explained as it was the only concept used and that is the reason why it is emphasized more. Also the identity of Supportive appears prominent which means that they helped more each other in order to solve the exercise.

3.3.5.3 Comparison of the Groups that used the fundamentals Concepts of Abstract and Inheritance and the Groups that did not use any of the fundamentals Concepts

In this case, the difference on the first and the second dimension is not significantly different::

- *Along the X axis, a two sample t test assuming unequal variance showed Abstract,Inheritance (mean=0.548, SD=0.693, N=18 was not statistically significantly different at the alpha=0.05 level from - (mean=1.166, SD=1.007, N=8; t(10.072)= 1.579, p=0.145, Cohen's d=0.776).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Abstract,Inheritance (mean=0.076, SD=1.397, N=18 was not statistically significantly different at the alpha=0.05 level from - (mean=-0.325, SD=1.439, N=8; t(13.141)= 0.661, p=0.520, Cohen's d=0.284).*

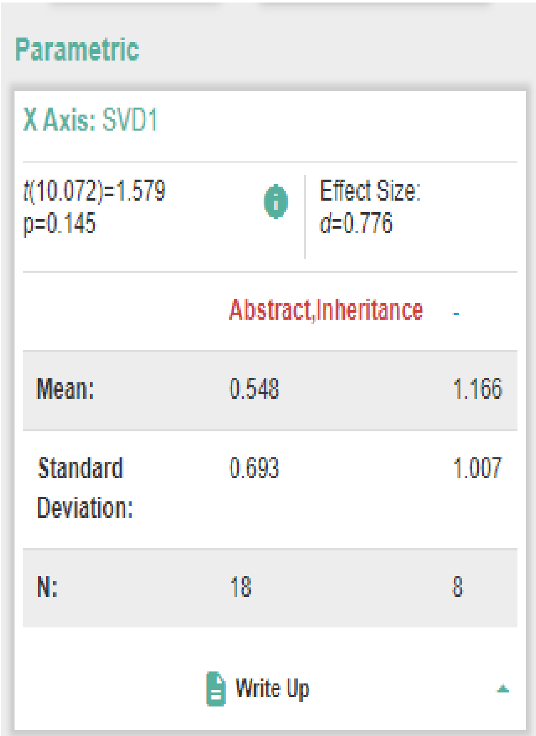


Figure 3-152 The results of the Independent T-Test for the first dimension

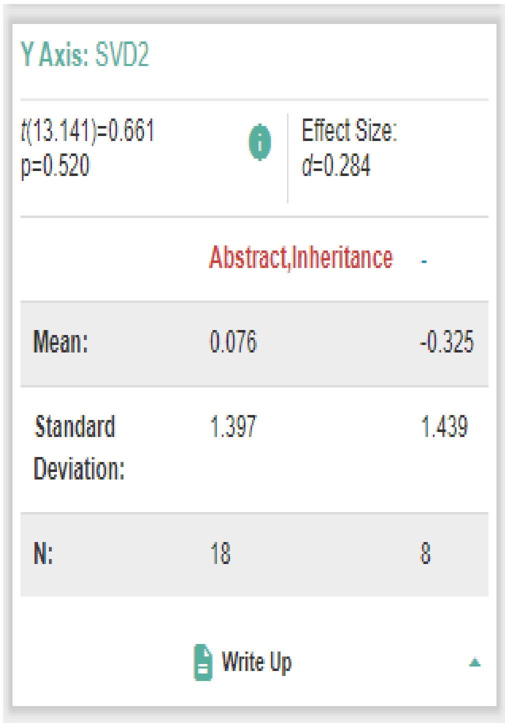
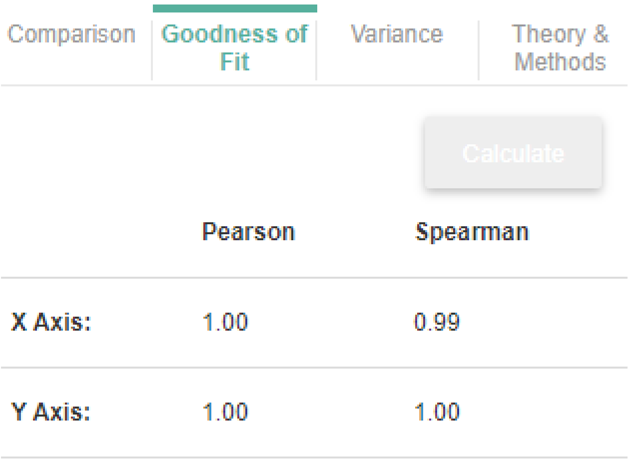


Figure 3-153 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-152 and Figure 4-153 show the independent t-test for the categories Abstract, Inheritance and None. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson’s and Spearman’s r. In this case, both are equal to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved.



	Pearson	Spearman
X Axis:	1.00	0.99
Y Axis:	1.00	1.00

Figure 3-154 Pearson’s and Spearman’s R for the Goodness of Fit

The Figure 4-154 shows the goodness of fit for the data with the statistics of Pearson’s and Spearman’s R. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two categories the two following networks for the two categories will appear. The Figure 4-155 shows the cognitive network for the None category and the Figure 4-156 shows the cognitive network for the Abstract, Inheritance category.

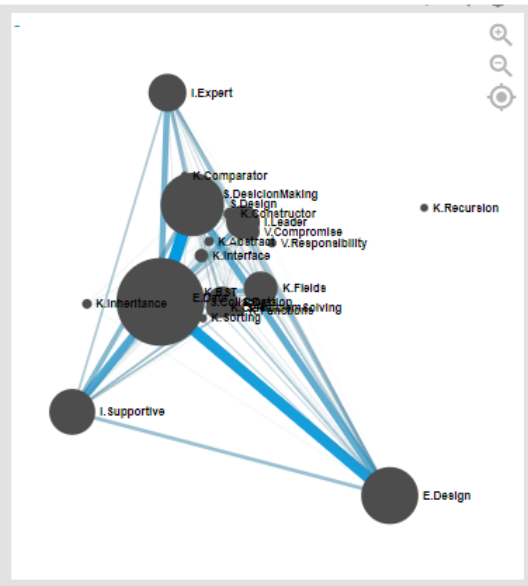


Figure 3-155 The Network of the centroid for None(-)

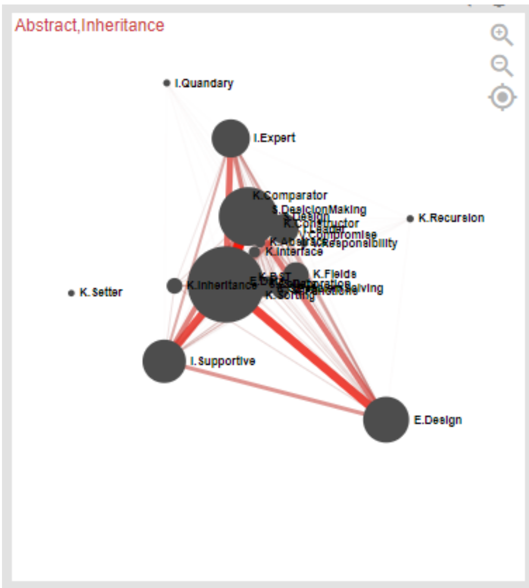


Figure 3-156 The Network of the centroid for Abstract, Inheritance

There are no major differences between the two networks analyzed. The differences worth presenting are:

The members of the groups that used the concepts of Abstract and Inheritance have more connections with the Identity of Supportive and Skill of Collaboration node. More specifically the stronger connections are formed between the Identity of Supportive and the Epistemology of Design (E.Design-I.Supportive: 0.04) the Knowledge of Inheritance (I.Supportive-K.Inheritance: 0.03) and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.046). There are also strong connections between the Skill of Collaboration and the Knowledge of Inheritance (K.Inheritance-S.Collaboration: 0.048) and Functions (K.Functions-S.Collaboration: 0.052) as well as the Identity of Expert (I.Expert-S.Collaboration: 0.053). Also there are strong connections between the Skill of Design and the Knowledge of Functions (K.Functions-S.Design: 0.042). That means that the members had design skills and the knowledge needed to solve the exercise, and they tend to confirm the proposition of one another more often.

The members of the groups that did not use any of the concepts have more connections starting from the Skill of Data. In particular there are strong connections between the Skill of Data and the Identity of Supportive (S.Data-I.Supportive: 0.04), the Knowledge of Fields (S.Data-K.Fields: 0.04) and the Epistemology Design (S.Data-E.Design: 0.04). There are also strong connections between the Skill of Collaboration and the Skill of Data (S.Collaboration-S.Data: 0.07) and the Epistemology of Design (S.Collaboration-E.Design: 0.03). Lastly another connection worth mentioning is the one between the Value of Compromise and the Skill of Design (S.Design-V.Compromise:0.03). That means that this category has better skills for handling the data than the design of code. Also the connections with the compromise value and the epistemology of design show that they offer and accept opinions and confirm the opinions of others.

3.3.5.4 Comparison of the Groups that used the fundamentals Concepts of Abstract, Inheritance and Comparator and the Groups that used the fundamentals Concept of Inheritance

In this case, the difference on the first and the second dimension is not significantly different::

- *Along the X axis, a two sample t test assuming unequal variance showed Abstract,Inheritance,Comparator (mean=-0.397, SD=0.812, N=7 was not statistically significantly different at the alpha=0.05 level from Inheritance (mean=-2.735, SD=2.430, N=6; t(5.956)= -2.252, p=0.066, Cohen's d=1.341).*
- *Along the Y axis, a two sample t test assuming unequal variance showed Abstract,Inheritance,Comparator (mean=0.451, SD=1.628, N=7 was not statistically significantly different at the alpha=0.05 level from Inheritance (mean=-0.321, SD=1.786, N=6; t(10.304)= 0.808, p=0.437, Cohen's d=0.453).*

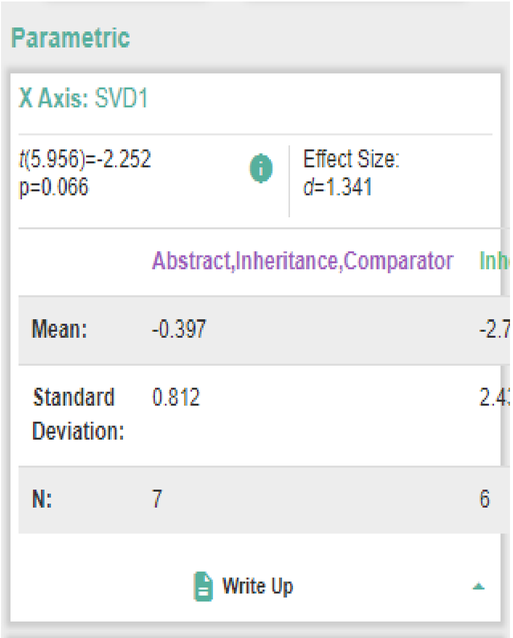


Figure 3-157 The results of the Independent T-Test for the first dimension

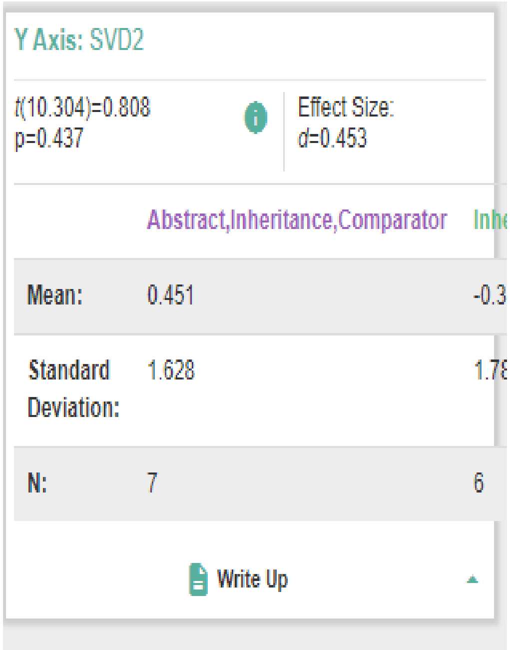
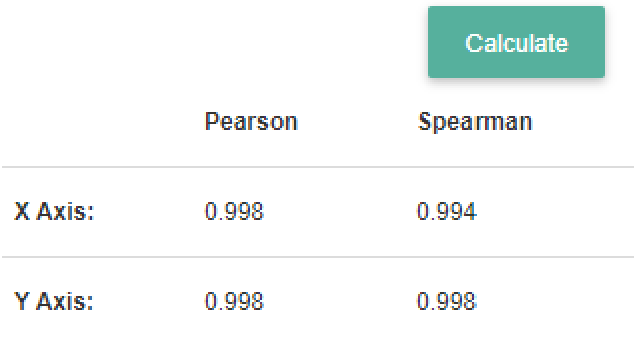


Figure 3-158 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-157 and Figure 4-158 show the independent t-test for the categories Abstract, Inheritance, Comparator and Inheritance. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are equal to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved.



	Pearson	Spearman
X Axis:	0.998	0.994
Y Axis:	0.998	0.998

Figure 3-159 Pearson's and Spearman's R for the Goodness of Fit

The Figure 4-159 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two categories the two following networks for the two categories will appear. The Figure

4-160 shows the cognitive network for the Inheritance category and the Figure 4-161 shows the cognitive network for the Abstract, Inheritance, Comparator category.

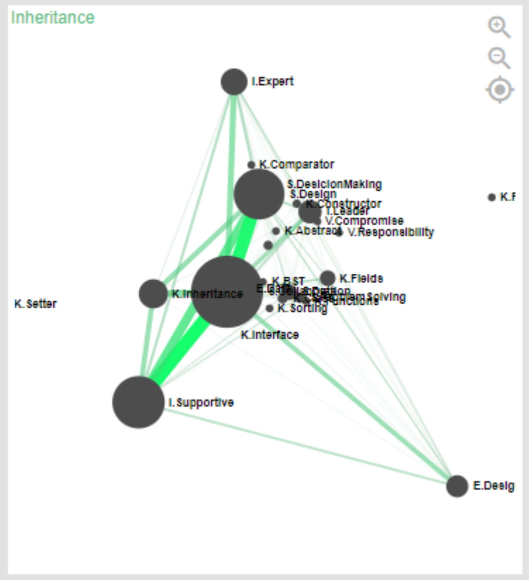


Figure 3-160 The Network of the centroid for Inheritance

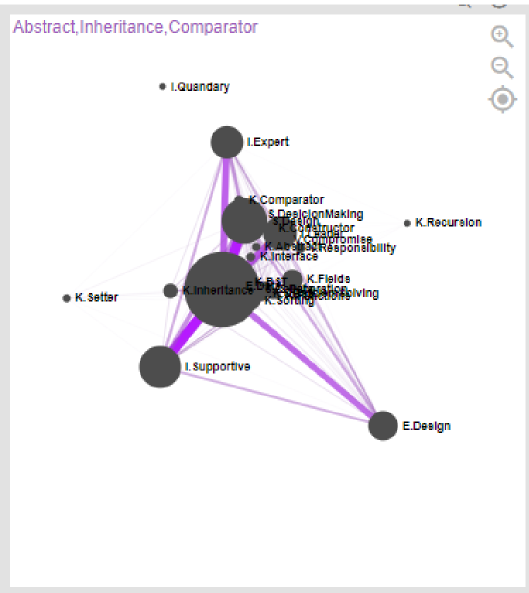


Figure 3-161 The Network of the centroid for Abstract, Inheritance, Comparator

There are no major differences between the two networks analyzed. The differences worth presenting are:

The members of the groups that used the concepts of Abstract, Inheritance and Comparator have more connections in comparison to the other category starting from the Epistemology of Design and the Skill of Collaboration. In particular there are strong connections between the Skill of Collaboration and the Identity of Expert (I.Expert-S.Collaboration: 0.078), the Identity of Leader (I.Leader-S.Collaboration: 0.155) and the Knowledge of Sorting (K.Sorting-S.Collaboration: 0.107). There are also strong connections between the Identities of Supportive and Leader (I.Supportive-I.Leader: 0.055) and between the Epistemology of Design and the Skills of Collaboration (S.Collaboration-E.Design: 0.089) and Design (S.Design-E.Design: 0.06). That means this category has distinctive roles, with the Leader and Expert Identities more prominent than the others. Also the collaboration, the epistemology and the skill of design have higher values which mean that this category has better developing and collaborative skills.

The members of the groups that used the concepts of Inheritance have more connections starting from the Identity of Supportive and the Knowledge of Inheritance. In particular there are strong connections between the Identity of Supportive and the Skill of Design (I.Supportive-S.Design: 0.124), the Skill of Collaboration (I.Supportive-S.Collaboration: 0.049) and the Knowledge of Inheritance (I.Supportive- K.Inheritance: 0.142). There are also strong connections between the Knowledge of Inheritance and the Skill of Design (K.Inheritance-S.Design: 0.130) and the Skill of Collaboration (K.Inheritance-S.Collaboration: 0.141). That means this category focused more on Inheritance which can be explained as it was the only concept used and that is the reason why it is emphasized more. Also the identity of Supportive appears prominent which means that they helped more each other in order to solve the exercise.

3.3.5.5 Comparison of the Groups that used the fundamentals Concepts of Abstract, Inheritance and Comparator and the Groups that did not use any of the fundamentals Concepts

In this case there is not a significant difference on the second dimension but there is a significant difference on the first one:

- *Along the X axis, a two sample t test assuming unequal variance showed - (mean=1.166, SD=1.007, N=8 was statistically significantly different at the alpha=0.05 level from Abstract,Inheritance,Comparator (mean=-0.397, SD=0.812, N=7; t(12.935)= -3.326, p=0.006, Cohen's d=1.696).*
- *Along the Y axis, a two sample t test assuming unequal variance showed - (mean=-0.325, SD=1.439, N=8 was not statistically significantly different at the alpha=0.05 level from Abstract,Inheritance,Comparator (mean=0.451, SD=1.628, N=7; t(12.142)= -0.971, p=0.350, Cohen's d=0.507).*

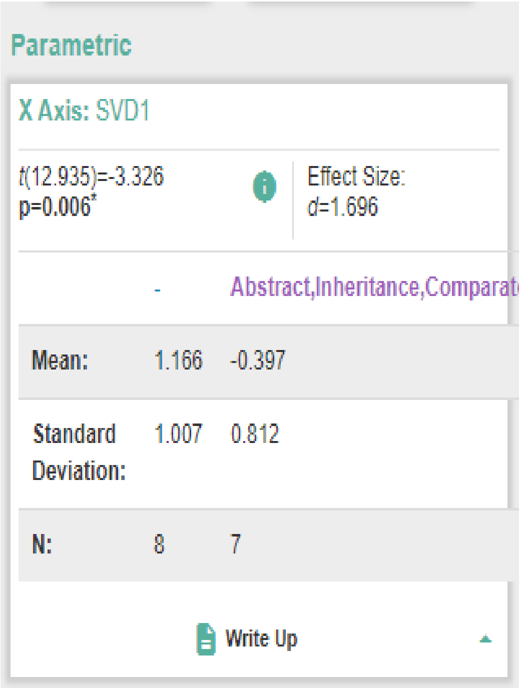


Figure 3-162 The results of the Independent T-Test for the first dimension

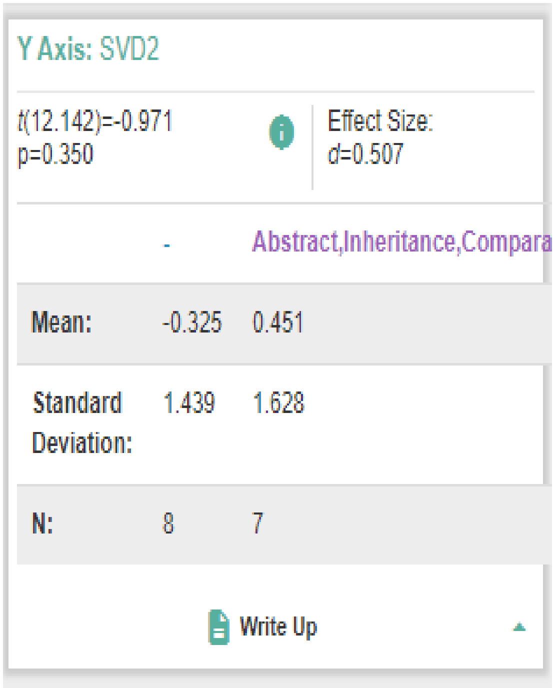


Figure 3-163 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-162 and Figure 4-163 show the independent t-test for the categories Abstract, Inheritance, Comparator and None. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are equal to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved.

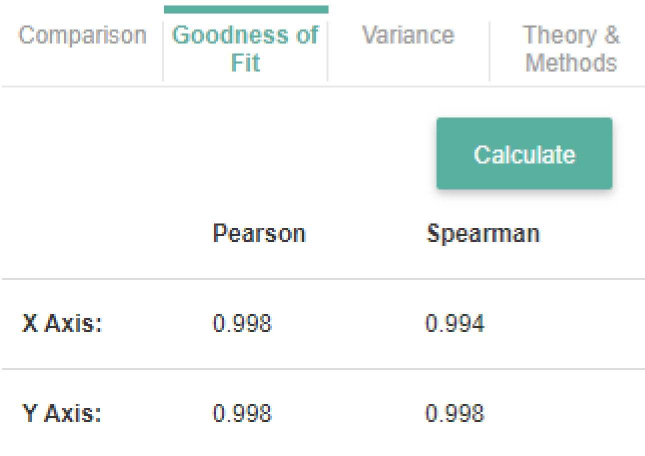


Figure 3-164 Pearson's and Spearman's R for the Goodness of Fit

The Figure 4-164 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these two categories the two following networks for the two categories will appear. The Figure 4-165 shows the cognitive network for the None category and the Figure 4-166 shows the cognitive network for the Abstract, Inheritance, Comparator category.

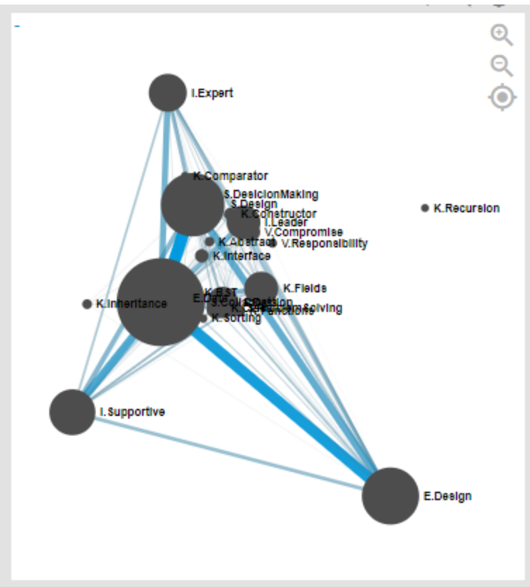


Figure 3-165 The Network of the centroid for None

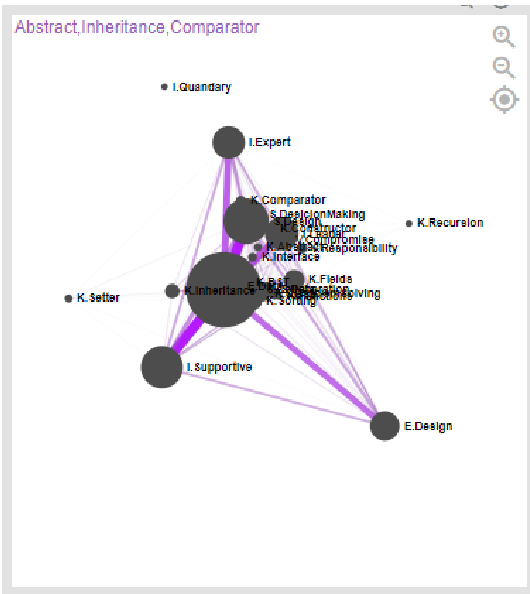


Figure 3-166 The Network of the centroid for Abstract, Inheritance, Comparator

There are many differences between the two networks analyzed. The differences worth presenting are:

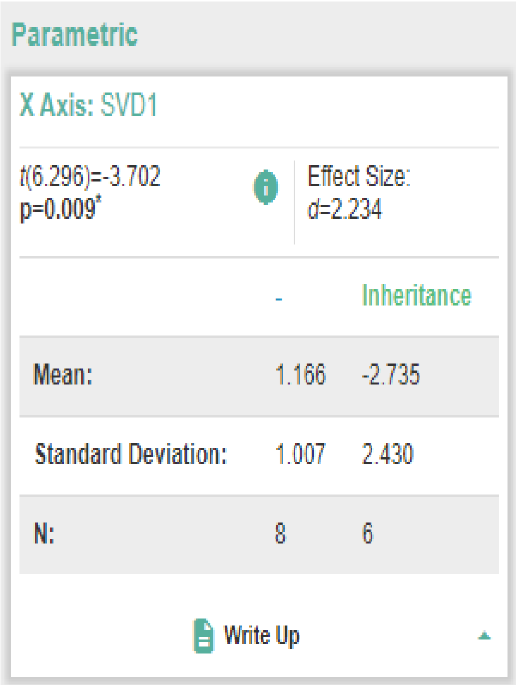
The members of the groups that used the concepts of Abstract, Inheritance and Comparator have more connections in comparison to the other category starting from the Identity of Supportive and the Skill of Collaboration. In particular there are strong connections between the Skill of Collaboration and the Knowledge of Inheritance (K.Inheritance-S.Collaboration: 0.064), the Identity of Leader (I.Leader-S.Collaboration: 0.124) and the Knowledge of Sorting (K.Sorting-S.Collaboration: 0.104). There are also strong connections between the Identities of Supportive and Leader (I.Supportive-I.Leader: 0.074) and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.128). The members of this category support each other but also we can see that the identity roles are prominent in this category.

The members of the groups that did not use any of the concepts have more connections starting from the Epistemology Design. In particular there are strong connections between the Epistemology Design and the Skill of Collaboration (E.Design-S.Collaboration: 0.126), the Knowledge of Fields (E.Design -K.Fields: 0.045) and the Identity of Expert (I.Expert-E.Design: 0.059). There are also strong connections between the Skill of Design and the Epistemology Design (S.Design-E.Design: 0.122) and between the Epistemology of Design and the Skill of Data (S.Data-E.Design: 0.066).The connections with the epistemology of design show that they offer opinions and confirm the opinions of others.

3.3.5.6 Comparison of the Groups that used the fundamentals Concept of Inheritance and the Groups that did not use any of the fundamentals Concepts

In this case there is not a significant difference on the second dimension but there is a significant difference on the first one:

- *Along the X axis, a two sample t test assuming unequal variance showed - (mean=1.166, SD=1.007, N=8 was statistically significantly different at the alpha=0.05 level from Inheritance (mean=-2.735, SD=2.430, N=6; t(6.296)= -3.702, p=0.009, Cohen's d=2.234).*
- *Along the Y axis, a two sample t test assuming unequal variance showed - (mean=-0.325, SD=1.439, N=8 was not statistically significantly different at the alpha=0.05 level from Inheritance (mean=-0.321, SD=1.786, N=6; t(9.453)= -0.005, p=0.996, Cohen's d=0.003).*



Parametric

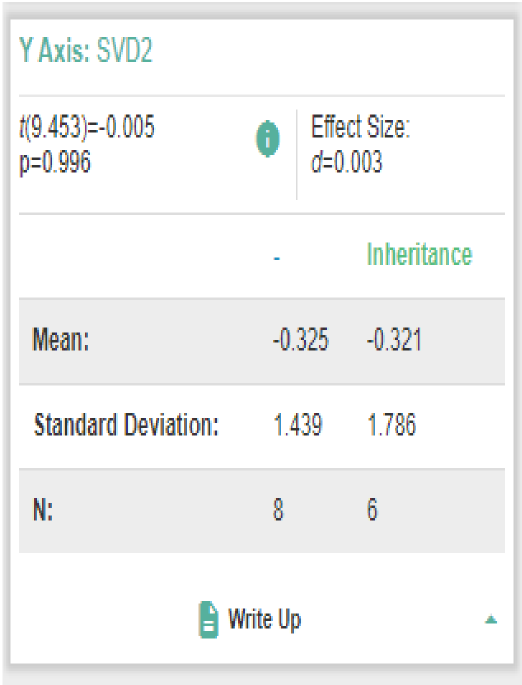
X Axis: SVD1

$t(6.296)=-3.702$ $p=0.009^*$ | Effect Size: $d=2.234$

	-	Inheritance
Mean:	1.166	-2.735
Standard Deviation:	1.007	2.430
N:	8	6

Write Up

Figure 3-167 The results of the Independent T-Test for the first dimension



Y Axis: SVD2

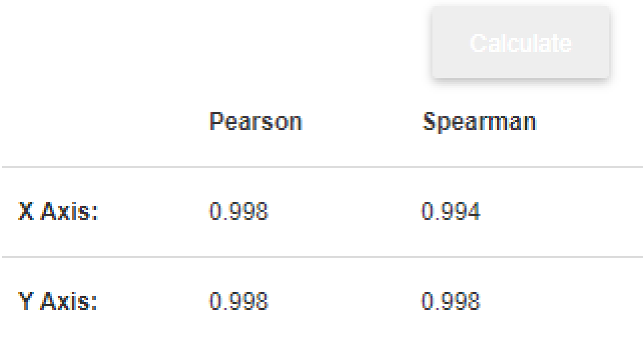
$t(9.453)=-0.005$ $p=0.996$ | Effect Size: $d=0.003$

	-	Inheritance
Mean:	-0.325	-0.321
Standard Deviation:	1.439	1.786
N:	8	6

Write Up

Figure 3-168 The results of the Independent T-Test for the second dimension

The figures above; Figure 4-167 and Figure 4-168 show the independent t-test for the categories Inheritance and None. Also the strength of the correlation between the centroids and the projected points in the model can be calculated using both Pearson's and Spearman's r. In this case, both are equal to 1 for both dimensions, because the number of units in the model is small compared to the number of dimensions. Optimization is therefore easy to be solved.



Calculate

	Pearson	Spearman
X Axis:	0.998	0.994
Y Axis:	0.998	0.998

Figure 3-169 Pearson's and Spearman's R for the Goodness of Fit

The Figure 4-169 shows the goodness of fit for the data with the statistics of Pearson's and Spearman's R. To determine the differences in link structures between the two networks we look at the equiload projections. Thus clicking on the centroids of these

two categories the two following networks for the two categories will appear. The Figure 4-170 shows the cognitive network for the None category and the Figure 4-171 shows the cognitive network for the Inheritance category

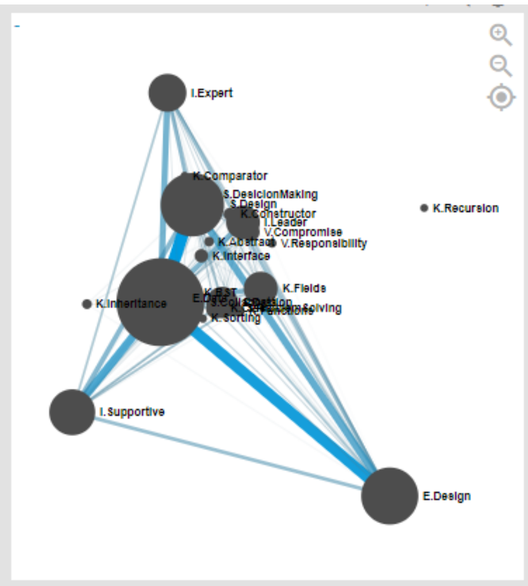


Figure 3-170 The Network of the centroid for None

Figure 3-171 The Network of the centroid for Inheritance

There are many differences between the two networks analyzed. The differences worth presenting are:

The members of the groups that did not use any of the concepts have more connections starting from the Epistemology Design and the Skill of Collaboration. In particular there are strong connections between the Epistemology Design and the Skill of Design (E.Design-S.Design: 0.182), the Knowledge of Fields (E.Design -K.Fields: 0.07) and the Identity of Expert (I.Expert-E.Design: 0.089). There are also strong connections between the Skill of Collaboration and the Epistemology Design (S.Design-E.Design: 0.216), the Knowledge of Fields (K.Fields-S.Collaboration: 0.091) and the Skill of Data (S.Collaboration –S.Data:0.066). The connections with the epistemology of design show that they collaborate more and mainly confirm the opinions of others.

The members of the groups that used the concepts of Inheritance have more connections starting from the Identity of Supportive and the Knowledge of Inheritance. In particular there are strong connections between the Identity of Supportive and the Skill of Design (I.Supportive-S.Design: 0.138), the Skill of Collaboration (I.Supportive-S.Collaboration: 0.178) and the Knowledge of Inheritance (I.Supportive- K.Inheritance: 0.17). There are also strong connections between the Knowledge of Inheritance and the Skill of Design (K.Inheritance-S.Design: 0.165) and the Skill of Collaboration

(K.Inheritance-S.Collaboration: 0.205). That means this category focused more on Inheritance which can be explained as it was the only concept used and that is the reason why it is emphasized more. Also the identity of Supportive appears prominent which means that they helped more each other in order to solve the exercise.

3.4 Summary

3.4.1 RQ1a. What types of connections between codes are made by each students' Group?

The stronger connection of the Group1 is between the Skill of Collaboration and the Skill of Design (S.Collaboration-S.Design: 0.439). There are also strong connections between the Identity of Expert and the Skills of Design (I.Expert-S.Design: 0.199) and Collaboration (I.Expert-S.Collaboration: 0.190). Lastly there are strong connections between the Epistemology of Design and the Skills of Collaboration (E.Design-S.Collaboration: 0.273) and Design (E.Design-S.Design: 0.292).

The stronger connection of Group2 is between the Skill of Collaboration and the Skill of Design (S.Collaboration-S.Design: 0.393). There are also strong connections between the Identity of Expert and the Skills of Design (I.Expert-S.Design: 0.135) and Collaboration (I.Expert-S.Collaboration: 0.182). Lastly there are strong connections between the Epistemology of Design and the Skills of Collaboration (E.Design-S.Collaboration: 0.373) and Design (E.Design-S.Design: 0.281).

The stronger connection of Group3 is between the Skill of Collaboration and the Skill of Design (S.Collaboration-S.Design: 0.394). There are also strong connections between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.260), between the Identity of Leader and the Skill of Collaboration (I.Leader-S.Collaboration: 0.258) and between the Identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.318). Lastly there are strong connections between the Epistemology of Design and the Skills of Collaboration (E.Design-S.Collaboration: 0.387) and Design (E.Design-S.Design: 0.216).

The stronger connection of Group4 is between the Skill of Collaboration and the Skill of Design (S.Collaboration-S.Design: 0.428). There are also strong connections between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.277), between the Identity of Supportive and the Skill of Design (I.Supportive-S.Design: 0.249) and between the Identity of Supportive and the Skill of Collaboration

(I.Supportive-S.Collaboration: 0.371). Lastly there are strong connections between the Epistemology of Design and the Skills of Collaboration (E.Design-S.Collaboration: 0.345) and Design (E.Design-S.Design: 0.241).

The stronger connection of Group5 is between the Skill of Collaboration and the Skill of Design (S.Collaboration-S.Design: 0.459). There are also strong connections between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.301), between the Identity of Supportive and the Skill of Design (I.Supportive - S.Design: 0.241) and between the Identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.362). Lastly there are strong connections between the Epistemology of Design and the Skills of Collaboration (E.Design-S.Collaboration: 0.231) and between the Identity of Leader and the Skill of Collaboration (I.Leader-S.Collaboration: 0.319).

The stronger connection of Group6 is between the Identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.471). There are also strong connections between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.228), between the Skill of Collaboration and the Skill of Design (S.Collaboration -S.Design: 0.437) and between the Epistemology of Design and the Skill of Collaboration (E.Design-S.Collaboration: 0.369). Lastly there are strong connections between the Identity of Supportive and the Skill of Design (I.Supportive-S.Design: 0.256)

The stronger connection of Group7 is between the Skill of Collaboration and the Skill of Design (S.Collaboration-S.Design: 0.447). There are also strong connections between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.332) and between the Identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.321). Lastly there are strong connections between the Epistemology of Design and the Skills of Collaboration (E.Design-S.Collaboration: 0.361) and Design (E.Design-S.Design: 0.207).

The stronger connection of Group8 is between the Skill of Collaboration and the Skill of Design (S.Collaboration-S.Design: 0.440). There are also strong connections between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.214), between the Identity of Supportive and the Skill of Design (I.Supportive - S.Design: 0.254) and between the Identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.388). Lastly there are strong connections between the

Epistemology of Design and the Skills of Collaboration (E.Design-S.Collaboration: 0.332) and Design (E.Design-S.Design: 0.217).

The stronger connection ofGroup9 is between the Identity of Supportive and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.458). There are also strong connections between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.244), between the Skill of Collaboration and the Skill of Design (S.Collaboration -S.Design: 0.323) and between the Epistemology of Design and the Skill of Collaboration (E.Design-S.Collaboration: 0.277). Lastly there are strong connections between the Identity of Leader and the Skill of Collaboration (I.Leader-S.Collaboration: 0.275).

The stronger connection of Group10 is between the Skill of Collaboration and the Identity of Supportive (S.Collaboration-I.Supportive: 0.475). There are also strong connections between the Identity of Expert and the Skill of Collaboration (I.Expert-S.Collaboration: 0.297), between the Identity of Supportive and the Skill of Design (I.Supportive -S.Design: 0.251) and between the Skill of Design and the Skill of Collaboration (I.Supportive-S.Collaboration: 0.441). Lastly there is strong connection between the Epistemology of Design and the Skills of Collaboration (E.Design-S.Collaboration: 0.242).

3.4.2 RQ1b. Is there a significant difference between the discourse networks of Group1 and the other nine different students' Groups?

In all the cases the statistical tests confirm that there are statistically significant differences between the Group1 and the other groups. The members of the Group 1 focused more on the Methods of the OOP because there are strong connections between Knowledge of Methods and the Skills of Design. The rest of the groups in general had stronger connections with the epistemology codes and the identity of Supportive.

3.4.3 RQ2a. What types of connections between codes are made by groups in the High-to-High category? What types of connections between codes are made by Groups in the High/Low-to-High category? What types of connections between codes are made by Groups in the High/Low-to-Low category?

The nodes that had the strongest connections in the High-to-High category network are S.Design, I.Expert, I.Supportive, S.Collaboration, and E.Design. Regarding the two Identity codes, the participants in this category had stronger connections with I.Expert because the members had more advanced knowledge and skills in computer programming and with the I.Supportive because this type of person helps to solve problems during the process. The connections with the Design codes indicate the students' abilities as a result of their performance in the OOP course. Overall, there were dense connections between the Knowledge Codes in this category.

The nodes that had the strongest connections in the High/Low-to-High category network are S.Design, I.Expert, I.Supportive, S.Collaboration, E.Design. The participants in this category had a strong connection with I.Expert code, which could be explained by the more advanced knowledge and skills in computer programming that some team members had. I.Supportive, seems to have stronger connections than the other Identity codes, which might mean that most of the time these students did not work off their own individual knowledge but relied on collaboration between team members. In this category, there are equally strong connections with/between E.Design and S.Design.All the connections with S.Design indicate the design abilities of students'. The connections with E.Design confirm the design ideas proposed.

The strongest connection in the category High/Low-to-Low is between E.Design and S.Collaboration. The participants had strong connections between I.Expert and I.Supportive, which can perhaps be explained by the fact that the majority of these students did not or were not able to propose any new ideas for code development. Instead, what they seemed to do is to often confirm the ideas proposed by the other members regardless of whether they were right or wrong. I.Supportive had much stronger connections than the other Identity codes, meaning that most of the time, these students did not work off their own individual knowledge but relied on collaboration between team members. In addition, in this category there were stronger connections with the E.Design than the S.Design code, indicating that these students did not have strong design abilities but preferred to simply confirm the proposals of other team members.

3.4.4 RQ2b. Is there a significant difference between the discourse networks of groups of the three categories: High-to-High, High/Low-to-High and High/Low-to-Low?

In the case of the High-to-High and High/Low-to-High the two categories are not significantly different. More specifically in the H-H category there are more connections between the knowledge codes and also skills like collaboration, design and data. The High/Low-to-High has less connections from the previous one but more pronounced. The most important ones are the supportive identity and the collaboration which means that all the members helped the other if needed the expert identity which means that the expert identity collaborated with the other members and finally the epistemology or confirmation of design and the collaboration which means that the members confirmed the design ideas of the other.

In the case of High-to-High and High/Low-to-Low the two categories are not significantly different. More specifically in the High-to-High category there are more connections between the knowledge codes and also skills like collaboration, design and data. The High/Low-to-Low has fewer connections from the previous one and the most important one is between the skill of collaboration and the epistemology of design. This connection can be explained because the members with less knowledge and skill usually just confirmed the others and do not propose anything new.

In the case of High/Low-to-High and High/Low-to-Low the two categories are again not statistically different. There are although differences between the two. In the High/Low-to-High category there are more connections between the knowledge codes, the user identities and also skills like collaboration and design. As for the High/Low-to-Low there are only 3 important connections to be analyzed. The first one is between the skill of data and the knowledge of fields. The second one is between the skill of data and collaboration and the third was between the skill of data and the epistemology-confirmation of design. This connection can be explained because the members with less knowledge and skill usually just confirmed the others and do not propose anything new. Also we can be observed that the skill of data be itself is not enough to solve a programming exercise without knowing programming design.

3.4.5 RQ3. Is there a significant difference between the discourse networks of groups of the same Category?

In the High-to-High category groups seem to be significantly different cause to the identity roles, the epistemology and the collaboration skills. The groups 5 and 8 have more pronounced the supportive identity which means that all the members helped the other if needed. The collaboration skill for the three groups is high along with the epistemology-confirmation of the design. The connection related to design means that the members of the groups were skillful in code development and also confirmed the design ideas of the other members.

The groups of the High/Low-to-High category have no big differences and their main characteristics are strong connection between the skills of collaboration, the epistemology-confirmation and the skill of design and the identities of supportive and the expert.

The groups of the High/Low-to-Low category seem to be significantly different cause to the skill of data, the epistemology of design and the supportive identity. More specifically the group 3 and the group 6 have more pronounced the supportive identity which means that all the members helped the other if needed. As for the group 2 the epistemology of design and the data handling skill are more profound. The results showed that the supportive identity was not enough by itself to ensure a good performance.

3.4.6 RQ4. Is there a significant difference between the discourse networks of the categories of the Groups based on the fundamental OOP concepts they used?

The groups of the category that used only the fundamentals concepts of Abstract and Inheritance in comparison to the ones that used all the concepts, have good skills and knowledge of the methods needed and they tend to confirm the proposition of one another more often. Also in comparison to the ones that used only the concept of Inheritance the members of this category had the design skills needed and they tend to confirm the proposition of one another more often. Lastly in comparison to the groups that did not use any of the concepts they had design skills and the knowledge needed to solve the exercise, and they tend to confirm the proposition of one another more often.

The groups of the category that used all the fundamentals concepts of OOP in comparison to the ones that used only the fundamentals of Abstract and Inheritance, have distinctive roles, better collaboration and knowledge of Sorting. Also the identity of Supportive appears prominent which means that they helped more each other in order to solve the exercise. In comparison to the ones that used only the concept of Inheritance, the members of this category had distinctive roles, with the Leader and Expert Identities more prominent than the others. Also the collaboration, the epistemology and the skill of design have higher values which mean that this category has better developing and collaborative skills. Lastly in comparison to the groups that did not use any of the concepts they support each other but also we can see that the identity roles are prominent in this category.

The groups of the category that used only the concept of Inheritance in comparison to the ones that used the concepts of Abstract and Inheritance, focused more on Inheritance which can be explained as it was the only concept used and that is the reason why it is emphasized more. Also the identity of Supportive appears prominent which means that they helped more each other in order to solve the exercise.In comparison to the groups that used all the fundamentals concepts of OOP they focused more on Inheritance which can be explained as it was the only concept used and that is the reason why it is emphasized more. Also the identity of Supportive appears prominent which means that they helped more each other in order to solve the assignment. Lastly in comparison to the groups that did not use any of the concepts they focused more on Inheritance which can be explained as it was the only concept used and that is the reason why it is emphasized more. Also the identity of Supportive appears prominent which means that they helped more each other in order to solve the exercise.

The groups of the category that did not use any of the concept in comparison to the groups that used only the concepts of Abstract and Inheritance had better skills for handling the data than the design of code. Also the connections with the compromise value and the epistemology of design show that they offer and accept opinions and confirm the opinions of others. In comparison to the groups that used all the fundamentals concepts of OOP they tend to offer opinions and confirm the opinions of others. Lastly in comparison to the groups that used only the concept of Inheritance they collaborate more and mainly confirm the opinions of others.

4 Conclusion

4.1 Summary and conclusions

As the educational data and the resources nowadays were increased, created the need to evaluate and analyze them but also to develop new more efficient techniques. With ENA we can analyze the knowledge, the skills but also the social behavior of the participants in educational activities. In this Thesis we presented many applications of ENA in different scientific fields, such as the Medical and Education fields and we also used ENA to analyze collaborative code development for the solution of an OOP assessment.

In this Thesis we presented the theoretical background of ENA. We conducted a Literature Review on ENA applications in different fields of Education for both students and teachers and also in the Medical field. Based on the findings of the Literature Review we carried out a research with the participation of 37 students of the Department of Applied Informatics, University of Macedonia, who worked collaboratively to solve an assignment using OOP. We proposed a coding scheme of OOP elements using Epistemic Frame Theory in order to analyze how students are collaborated using chat messages for solving an OOP assignment and thus shade light on what type of connections are made in the groups of students with different computer programming skills.

The results of the research showed that there was significant difference of the discourse networks of the majority of groups comparing with the discourse network of Group1, which members had high programming skills and high score in the collaborative OOP assignment. Also the comparison of the discourse networks of different categories of groups based on the students' course and assignment scores, showed that there was significant difference between all of them. However, the comparison of discourse networks of groups that belong to the same category of high-performance, or mixed-performance, or low-performance categories, for the most part did not show significant differences. Regarding the comparison of the discourse networks of the different categories of groups, based on the use of fundamental concepts of OOP, the differences between them were significant.

Through the research, we came to several conclusions regarding the behavior and performance of students during collaborative programming. In particular, we concluded that although the participants who belonged to the same category their behavior differed during the solution of the assignment. In some cases, as in the case of the category of groups High-to-High, Group 1 has as its apparent identities of Leader and Expert. The other two groups in this category (Group5 and Group8) have the most prevalent the identity of Supportive. This practically means that in one case we have participants with high both programming and leadership skills, while in the other case the participants have the previous skills to a lesser extent but they are more cooperative. The solution process is therefore based in the first case on the instructions of individuals with knowledge of the expert and the leadership tendencies of either themselves or others while in the other case they rely mainly on cooperation and support material offered by group members. These observations can be also observed to the other groups that have similar characteristics to the ones described above concerning the other two group categories (High/Low-to-Low and High/Low-to-High).

There were also cases in which there was one participant in a group that did the most of the job done. In this case this student was the one that solved the exercise for the others by providing ready for use code, comments and guidance towards the other members of his group. There are two groups that had this specific characteristic. In this case the student that provided this information had more connections with both the Identities of Leader and Expert but also with the Skills of Collaboration and Design. The other members of his group had more prominent the Identity of Supportive and the connections with the Skill of Collaboration and the Epistemology Codes.

Regarding the categorization of the groups based on the basic concepts of the OOP, it was observed that each category had a different number of groups. More specifically 2 groups used all the fundamental concepts of OOP (Group5 and Group9), 5 groups used the fundamentals concepts Abstract and Inheritance (Group1, Group3, Group4, Group7 and Group8), One group used only the concept of Inheritance (Group10) and 2 groups did not use any of the fundamental concepts of OOP (Group2 and Group6). Through dialogues it is found that in most cases the groups that did not use Comparator but used the other two concepts had the knowledge but did not have the time to develop a code for the issue that concerned it.

The results suggest that ENA provides information and new ideas for processes related to collaborative interactions in an OOP assignment.

4.2 Research limitations

The design of the current study is subject to limitations that concern the research part of the study. The findings of this study have to be seen in light of three main limitations. The first one is that the data collected and analyzed were only text messages. That basically means that there were concepts of analysis that if they are not mentioned in texts we could not know if they occur or not. Secondly, the nature of the research means that any conclusions drawn from the findings regarding students' abilities are limited to the study sample which participated in collaboratively solving the specific OOP assignment in the particular Computer Science course. Thirdly, the sample size (37) of the participants may have led to non-significant results during comparative analysis. Further studies will be conducted, which will include a larger sample size and with a higher number and greater variety of OOP assignments.

4.3 Future extensions

A future development of this work could be the analysis of the second exercise and the integration into the already existing data so that additional conclusions can be drawn. A future research topic could be a more rigorous analysis of the dialogues based on both the literal interpretation given and the actual interpretation performed by experts in order to avoid the risk of prejudice and for the result to be unbiased. Another further development of the research could be the application of other network analysis methods to extract additional data related to the network and the connections that exist in it.

References

Andrist, S. (2016). *Gaze Mechanisms for Situated Interaction with Embodied Agents*. The University of Wisconsin-Madison.

Andrist, S., Collier, W., Gleicher, M., Mutlu, B., & Shaffer, D. (2015). Look together: Analyzing gaze coordination with epistemic network analysis. Frontiers in psychology, 6, 1016.

Andrist, S., Ruis, A. R., & Shaffer, D. W. (2018). A network analytic approach to gaze coordination during a collaborative task. Computers in Human Behavior, 89, 339-348.

Arastoopour, G., & Shaffer, D. W. (2013). Measuring social identity development in epistemic games.

Arastoopour, G., Chesler, N. C., D'Angelo, C. M., Shaffer, D. W., Opgenorth, J. W., Reardan, C. B., Haggerty, N. P., & Lepak, C. G. (2012, June). Nephrotex: Measuring first-year students' ways of professional thinking in a virtual internship. In 2012 ASEE Annual Conference & Exposition (pp. 25-971).

Arastoopour, G., Chesler, N. C., Shaffer, D. W., & Swiecki, Z. (2015, June). Epistemic network analysis as a tool for engineering design assessment. In 2015 ASEE Annual Conference & Exposition (pp. 26-679).

Arastoopour, G., Shaffer, D. W., Swiecki, Z., Ruis, A. R., & Chesler, N. C. (2016). Teaching and assessing engineering design thinking with virtual internships and epistemic network analysis. International Journal of Engineering Education, 32(2), 1492-1501.

Baheti, P., Gehringer, E., & Stotts, D. (2002, August). Exploring the efficacy of distributed pair programming. In Conference on Extreme Programming and Agile Methods (pp. 208-220). Springer, Berlin, Heidelberg.

Bressler, D. M., Bodzin, A. M., Eagan, B., & Tabatabai, S. (2019). Using epistemic network analysis to examine discourse and scientific practice during a collaborative game. Journal of Science Education and Technology, 28(5), 553-566.

Chesler, N. C., Arastoopour Irgens, G., D'angelo, C. M., Bagley, E. A., & Shaffer, D. W. (2013). Design of a professional practice simulator for educating and motivating first-year engineering students. Advances in Engineering Education.

Chesler, N. C., Ruis, A. R., Collier, W., Swiecki, Z., Arastoopour, G., & Williamson Shaffer, D. (2015). A novel paradigm for engineering education: Virtual internships with individualized mentoring and assessment of engineering thinking. Journal of biomechanical engineering, 137(2), 024701.

Cockburn, A., & Williams, L. (2000). The costs and benefits of pair programming. Extreme programming examined, 8, 223-247.

Collier, W., Ruis, A., & Shaffer, D. W. (2016). Local versus global connection making in discourse. Singapore: International Society of the Learning Sciences.

Csanadi, A., Eagan, B., Shaffer, D., Kollar, I., & Fischer, F. (2017). Collaborative and individual scientific reasoning of pre-service teachers: New insights through epistemic network analysis (ENA). Philadelphia, PA: International Society of the Learning Sciences..

Curricula Recommendations, Accessed 9August, Retrieved from https://www.acm.org/education/curricula-recommendations, 2021

Dimeo, M., & Ruis, A. Thinking about Sources as Data: Reflections on Epistemic Network Analysis as Technique for Historical Research.

Epistemic Network Analysis, Accessed 9 August, Retrieved fromhttps://www.epistemicnetwork.org/, 2021.

Fisher, K. Q., Hirshfield, L., Siebert-Evenstone, A., Irgens, G. A., & Koretsky, M. (2016, June). Network analysis of interactions between students and an instructor during design meetings. In 2016 ASEE Annual Conference & Exposition.

Fougt, S. S., Siebert-Evenstone, A., Eagan, B., Tabatabai, S., & Misfeldt, M. (2018, March). Epistemic network analysis of students' longer written assignments as formative/summative evaluation. In Proceedings of the 8th International Conference on Learning Analytics and Knowledge (pp. 126-130).

Hatfield, D. (2015). The right kind of telling: An analysis of feedback and learning in a journalism epistemic game. International Journal of Gaming and Computer-Mediated Simulations (IJGCMS), 7(2), 1-23.

Herder, T., Swiecki, Z., Fougt, S. S., Tamborg, A. L., Allsopp, B. B., Shaffer, D. W., & Misfeldt, M. (2018, March). Supporting teachers' intervention in students' virtual collaboration using a network based model. In Proceedings of the 8th International Conference on Learning Analytics and Knowledge (pp. 21-25).

Johnson, D. W., Johnson, R. T., & Taylor, B. (1993). Impact of cooperative and individualistic learning on high-ability students' achievement, self-esteem, and social acceptance. The Journal of Social Psychology, 133(6), 839-844.

Knight, S., Arastoopour, G., Williamson Shaffer, D., Buckingham Shum, S., & Littleton, K. (2014). Epistemic networks for epistemic commitments.

Lund, K., Quignard, M., & Shaffer, D. W. (2017). Gaining Insight by Transforming between Temporal Representations of Human Interaction. Journal of Learning Analytics, 4(3), 102-122.

Mochizuki, T., Sasaki, H., Yamaguchi, Y., Hirayama, R., Kubota, Y., Eagan, B., Yuki, N., Funaoi, H., Kubota, Y., Suzuki, H., & Kato, H. (2019(b), October). Effects of Perspective-Taking Through Tangible Puppetry in Microteaching and Reflection on the Role-Play with 3D Animation. In International Conference on Quantitative Ethnography (pp. 315-325). Springer, Cham.

Nash, P., & Shaffer, D. W. (2013). Epistemic trajectories: Mentoring in a game design practicum. Instructional Science, 41(4), 745-771.

Orrill, C. H., & Shaffer, D. W. (2012). Exploring connectedness: Applying ENA to teacher knowledge.

Orrill, C., Shaffer, D. W., & Burke, J. (2013, April). Exploring coherence in teacher knowledge using epistemic network analysis. In American Educational Research Association Annual Conference (AERA 2013) (Vol. 27).

Rolim, V., Ferreira, R., Lins, R. D., & Găsević, D. (2019). A network-based analytic approach to uncovering the relationship between social and cognitive presences in communities of inquiry. The Internet and Higher Education, 42, 53-65.

Ruis, A. R. (2016). Trois Empreintes d'un Même Cachet": Toward a Historical Definition of Nutrition. Viral Networks: Connecting Digital Humanities and Medical History, 179-212.

Ruis, A. R., Rosser, A. A., Nathwani, J. N., Beems, M. V., Jung, S. A., & Pugh, C. M. (2019, October). Multiple uses for procedural simulators in continuing medical education contexts. In International Conference on Quantitative Ethnography (pp. 211-222). Springer, Cham.

Ruis, A. R., Rosser, A. A., Quandt-Walle, C., Nathwani, J. N., Shaffer, D. W., & Pugh, C. M. (2018). The hands and head of a surgeon: Modeling operative competency with multimodal epistemic network analysis. The American Journal of Surgery, 216(5), 835-840.

Ruis, A., Siebert-Evenstone, A., Pozen, R., Eagan, B. R., & Shaffer, D. W. (2018). A method for determining the extent of recent temporal context in analyses of complex, collaborative thinking. International Society of the Learning Sciences, Inc.[ISLS]..

Rupp, A. A., Choi, Y., Gushta, M., Mislevy, R., Bagley, E., Nash, P., Hatfield, D., Svarovsky, G., & Shaffer, D. (2009, June). Modeling learning progressions in epistemic games with epistemic network analysis: Principles for data analysis and generation. In Proceedings from the learning progressions in science conference (pp. 24-26). ISO 690

Schümmer, T., & Lukosch, S. G. (2009). Understanding tools and practices for distributed pair programming. Journal of Universal Computer Science, 15 (16), 2009.

Shaffer, D. (2018(b)). Big data for thick description of deep learning. Deep learning: Multi-disciplinary approaches.

Shaffer, D. W. (2004, October). Epistemic frames and islands of expertise: Learning from infusion experiences. In Embracing Diversity in the Learning Sciences: Proceedings of the Sixth International Conference of the Learning Sciences (p. 474). Psychology Press.

Shaffer, D. W. (2006). Epistemic frames for epistemic games. Computers & education, 46(3), 223-234.

Shaffer, D. W. (2018(a)). Epistemic network analysis: Understanding learning by using big data for thick description. International handbook of the learning sciences, 520-531.

Shaffer, D. W., & Ruis, A. R. (2017). Epistemic network analysis: A worked example of theory-based learning analytics. In C. Lang, G. Siemens, A. Wise, & D. Grasevic (Eds.), Handbook of Learning Analytics (pp. 175–187). Society for Learning Analytics Research.

Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. Journal of Learning Analytics, 3(3), 9-45.

Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., Frank, K, Rupp A.,& Mislevy, R. (2009). Epistemic network analysis: A prototype for 21st-century assessment of learning. International Journal of Learning and Media, 1(2).

Shaffer, D., & Ruis, A. (2017). Epistemic network analysis: A worked example of theory-based learning analytics. Handbook of learning analytics.
Shaffer, D., & Ruis, A. (2017). Epistemic network analysis: A worked example of theory-based learning analytics. Handbook of learning analytics.
Siebert-Evenstone, A. L., Irgens, G. A., Collier, W., Swiecki, Z., Ruis, A. R., & Shaffer, D. W. (2017(a)). In search of conversational grain size: modelling semantic structure using moving stanza windows. Journal of Learning Analytics, 4(3), 123-139.

Siebert-Evenstone, A. L., Irgens, G. A., Collier, W., Swiecki, Z., Ruis, A. R., & Shaffer, D. W. (2017(b)). In search of conversational grain size: modelling semantic structure using moving stanza windows. Journal of Learning Analytics, 4(3), 123-139.

Siebert-Evenstone, A., & Shaffer, D. W. (2019, October). Cause and Because: Using Epistemic Network Analysis to Model Causality in the Next Generation Science Standards. In International Conference on Quantitative Ethnography (pp. 223-233). Springer, Cham.

Silva L, A. J. Mendes, A. Gomes, "Computer-supported Collaborative Learning in Programming Education: A Systematic Literature Review.", 2020 IEEE Global Engineering Education Conference (EDUCON), pp. 1086-1095, 2020.

Sullivan, S., Warner-Hillard, C., Eagan, B., Thompson, R. J., Ruis, A. R., Haines, K., Pugh C.M. , Shaffer D.W., & Jung, H. S. (2018). Using epistemic network analysis to identify targets for educational interventions in trauma team communication. Surgery, 163(4), 938-943.

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.

Swiecki, Z., & Shaffer, D. W. (2018). Toward a taxonomy of team performance visualization tools. International Society of the Learning Sciences, Inc.[ISLS].

Wakimoto, T., Sasaki, H., Hirayama, R., Mochizuki, T., Eagan, B., Yuki, N., Funaoi, H., Kubota, Y., Suzuki, H., & Kato, H. (2019(a), October). Student teachers' discourse during puppetry-based microteaching. In International Conference on Quantitative Ethnography (pp. 234-244). Springer, Cham.

Williams, L., McDowell, C., Nagappan, N., Fernald, J., & Werner, L. (2003, October). Building pair programming knowledge through a family of experiments. In 2003 International Symposium on Empirical Software Engineering, 2003. ISESE 2003. Proceedings. (pp. 143-152). IEEE.

Wooldridge, A. R., Carayon, P., Shaffer, D. W., & Eagan, B. (2018). Quantifying the qualitative with epistemic network analysis: A human factors case study of task-allocation communication in a primary care team. IISE transactions on healthcare systems engineering, 8(1), 72-82.

Xinogalos, S., Satratzemi, M., Chatzigeorgiou, A., & Tsompanoudi, D. (2019). Factors affecting students' performance in distributed pair programming. Journal of Educational Computing Research, 57(2), 513-544.

Zhang, S., Liu, Q., & Cai, Z. (2019). Exploring primary school teachers' technological pedagogical content knowledge (TPACK) in online collaborative discourse: An epistemic network analysis. British Journal of Educational Technology, 50(6), 3437-3455.

Annex A – Paper presentation at the ICERI2021 conference

The following paper has been accepted for presentation at the ICERI2021 (14th annual International Conference of Education, Research and Innovation) conference and it'll be included in ICERI2021 proceedings.

Aliki Christou Maya Satratzemi, "EPISTEMIC NETWORK ANALYSIS OF STUDENTS' CHAT DATA ON A COLLABORATIVE SOLVING OF OBJECT-ORIENTED PROGRAMMING ASSIGNMENT", Proceedings of the14th annual International Conference of Education, Research and Innovation (ICERI2021), 8-9 November 2021 (https://iated.org/publications)