INVESTIGATING MATHEMATICS TEACHER EDUCATORS' SUBJECT MATTER KNOWLEDGE FOR TEACHING INFORMAL LINE OF BEST FIT

by

Jale Günbak

B.S., Secondary School Science and Mathematics Education, 2014

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Secondary School Science and Mathematics Education

Boğaziçi University

2019

INVESTIGATING MATHEMATICS TEACHER EDUCATORS' SUBJECT MATTER KNOWLEDGE FOR TEACHING INFORMAL LINE OF BEST FIT

APPROVED BY:

Assoc. Prof. Gülseren Karagöz Akar .....

(Thesis Supervisor)

Assoc. Prof. Serkan Özel .....

Assist. Prof. Sibel Kazak .....

DATE OF APPROVAL: 27.05.2019

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest gratitude to my supervisor Assoc. Prof. Gülseren Karagöz Akar who guided me throughout the entire process by offering her continuous advices and encouragements. It was privilege for me to work with her and to be benefited from her wisdom. I thank her for endless support, patience and the effort she put into guiding me during this study.

I am very grateful to my thesis committee members; Assoc. Prof. Serkan Özel and Assist. Prof. Sibel Kazak for their unconditional support and invaluable recommendations throughout the process.

I would like to thank Assoc. Prof. Fatma Aslan- Tutak and Assist. Prof. Yeşim İmamoğlu for their support and advices on my study.

I am very thankful to Merve Nur Belin for her invaluable support during the data collection process and for her help whenever I need.

I also would like to thank my colleagues Hanife Can Şen, Nisa Başara Baydilek, Betül Altay and Burcu Altun for their friendship, sincerity and encouragement. I always feel lucky to be a member of our room.

My special thanks go to Yücel Hatıl for always believing in me and being besides me whenever I need. I would never complete this study without your love and continuous encouragement. I am very grateful to share all beautiful feelings and moments with you. Love you!

I also owe special thanks to my father Hasan and my mother Şenay for always being besides me and believing in me. Their love and endless support always motivate me in every part of my life. I also would like to thank to my brothers Ahmet Emre and Gökay; and my sister Hanım. I am very lucky to have such a family. Gülse, you are sunshine of our family with your beautiful heart and mind.

ABSTRACT

INVESTIGATING MATHEMATICS TEACHER EDUCATORS' SUBJECT MATTER KNOWLEDGE FOR TEACHING INFORMAL LINE OF BEST FIT

In this study, I investigated eleven mathematics teacher educators' subject matter knowledge for teaching regarding the informal line of best fit. I specifically examined what conceptions and placement criteria for the informal line of best fit teacher educators had and what and how they used slope conceptualizations regarding the informal line of best fit. Basic qualitative research design was used in this study. I conducted approximately one hour long task-based interviews to understand teacher educators' conceptualizations. Results showed that five of the teacher educators had *representer*, four had *predictor* and two had *signal* as their dominant conceptions. Particularly, while teacher educators having the representer conception interpreted the informal line of best fit visually focusing on the data points particular to the situation at hand, the ones having the predictor conception additionally considered the predicted values. Further, the teacher educators having the signal conception focused on the data points particular to the situation, the predicted values on the best fit line and the residuals simultaneously and holistically. These results implied for the informal line of best fit that *signal conception* reflect and include the ideas of both *representer* and *predictor conceptions*. Results also indicated that seven teacher educators used *closest*, five of them referred to the *sum deviation* and two teacher educators also used *ellipse & area criterion* dominantly for placement of informal line of best fit. In addition, results regarding the slope conceptualizations, all teacher educators had *behaviour indicator* and *real world situation conceptions* of slope. Besides these conceptions of slope, participants holding representer conception had also *linear constant* and *parametric coefficient*, participants with predictor conception also had the *functional property*; and, participants having the signal conception were also able to interpret the slope as the *functional property*, *algebraic ratio* and *parametric coefficient* dominantly. Results altogether implied that although having all the aforementioned dominant conceptions are necessarily important to understand the reasoning behind the line of best fit, signal conception seems to be vital for teachers and teacher educators to have.

ÖZET

MATEMATİK ÖĞRETMEN EĞİTİMCİLERİN INFORMEL EN İYİ UYUM DOĞRUSUNU ÖĞRETMEK İÇİN KONU BİLGİSİNİN İNCELENMESİ

Bu çalışma 11 matematik öğretmen eğitimcisinin informel en iyi uyum doğru kavramını öğretmek için konu alan bilgisini incelemektir. Bu amaçla, öğretmen eğitimcilerinin informel en iyi uyum doğrusu için hangi kavramlara sahip olduğu, bu doğruyu yerleştirirken hangi kriterleri kullandığı ve bu konuya ilişkin hangi eğim kavramlarını nasıl kullandıkları incelenmiştir. Bu çalışmada temel nitel araştırma deseni kullanılmıştır. Öğretmen eğitimcilerin bu konuyu nasıl kavramlaştırdıklarını anlamak amacıyla, katılımcılarla yaklaşık 1 saat süren görüşmeler yapılmıştır. Sonuçlara göre, beş öğretmen eğitimci *'temsil eden'*, dört tanesi *'yordayıcı'* ve ikisinin de *'signal'* kavramını baskın olarak kullandığı görülmüştür. Ayrıca, *temsil eden* kavramına sahip olan öğretmen eğitimciler informel en iyi uyum doğrusunu eldeki verilere dayanarak görsel olarak yorumlarken, *yordayıcı* kavramına sahip olanlar ek olarak tahmin değerlerini de dikkate aldılar. Bunun yanısıra, *signal* kavramına sahip olan öğretmen eğitimciler eldeki verilere odaklanırken aynı zamanda tahmin değerlerini ve hata paylarını da bir bütün olarak ele aldılar. Bu sonuçlara göre, informel en iyi uyum doğrusuna yönelik olarak *signal* kavramının *temsil eden* ve *yordayıcı* kavramlarına ait fikirleri de içerdiği düşünülmektedir. Sonuçlar ayrıca doğru yerleştirirken yedi öğretmen eğitimcinin *'en yakın'*, beş tanesinin *'toplam deviyasyon'* ve ikisinin de *'elips & alan'* kriterini baskın olarak kullandığını göstermektedir. Son olarak, eğim kavramlarına ilişkin sonuçlara göre, bütün öğretmen eğitimciler *'davranış gösteren'* ve *'gerçek yaşam durumu'* kavramlarını kullanmıştır. Bunlara ek olarak, temsil eden kavramına sahip olan katılımcılar *'doğrusal sabit'* ve *'parametrik katsayı'*, yordayıcı kavramına sahip olan katılımcılar *'fonksiyonel özellik'* ve signal kavramına sahip olan katılımcılar ayrıca *'fonksiyonel özellik'*, *'cebirsel oran'* ve *'parametrik katsayı'* eğim kavramlarını baskın olarak kullanmıştır. Sonuçlar, en iyi uyum doğru kavramının temelinde yatan nedenleri anlamak için, yukarıda belirtilen bütün kavramların mutlaka önemli olduğunu göstermekle birlikte, öğretmenler ve öğretmen eğitimcilerinin signal kavramına sahip olmasının hayati önem taşıdığını göstermektedir.

TABLE OF CONTENTS

ACKNOWLEDGEMENT iii
ABSTRACT v
ÖZET vi
LIST OF TABLES ix
LIST OF ACRONYMS / ABBREVIATIONS xiii
1. INTRODUCTION 1
2. REVIEW OF LITERATURE 4
2.1. The Need for Studying Informal Line of Best Fit 4
2.2. Importance of Mathematics Teacher Educators' Knowledge 9
2.3. Studies on Conceptualization of Informal Line of Best Fit 11
2.3.1. Students' and Teachers' Conceptions of Informal Line of Best Fit 12
2.3.2. Criteria and Methods for Placement of Informal Line of Best Fit 13
2.3.3. Accuracy of Placements 15
2.3.4. Interpretation of Informal Line of Best Fit 16
2.4. Slope Conceptualization and Line of Best Fit 17
3. THEORETICAL FRAMEWORK 23
3.1. Statistical Knowledge For Teaching 23
3.2. Teachers' Conceptions and Criteria Framework for Informal Line of Best Fit 25
3.3. Slope Conceptualization Framework 28
4. METHOD 32
4.1. Research Design: Basic Qualitative Research 32
4.2. Participants 32
4.3. Data Collection 37

4.3.1. Interview Protocol 37
4.4. Data Analysis 42
5. RESULTS 50
5.1. Conceptions 50
5.2. Criteria and Methods 70
5.3. Placing a Line for Data with Nonassociation 92
5.4. Slope Conceptualization and Line of Best Fit 106
6. DISCUSSION AND CONCLUSION 119
6.1. Conceptions and Criteria 120
6.2. Slope Conceptualizations for Informal Line of Best Fit 126
7. IMPLICATIONS FOR RESEARCH AND TEACHING 132
REFERENCES 135
APPENDIX A: INTERVIEW PROTOCOL 142

LIST OF TABLES

Table 2.1. Conceptualizations of slope (Moore-Russo et al., 2010, p.9). 18

Table 3.1. Data sources used to investigate teacher educators' subject matter knowledge for teaching informal line of best fit. 24

Table 3.2. Conception categories (Casey, 2015b, p.17). 26

Table 3.3. Criteria categories (Casey, 2015b, p.17). 27

Table 3.4. Conceptualizations of slope (Moore-Russo et al., 2010, p.9). 29

Table 4.1. Participant information about their teaching and learning experiences. 34

Table 4.2. Summary of Tasks 1-7. 40

Table 4.3. Conception Framework (Casey, 2015b, p.17). 44

Table 4.4. Criteria Framework adopted and modified from Casey (2015b, p.17). 45

Table 5.1. Teacher educators' use of different conceptions including operational definitions, illustrative excerpts and frequency. 51

Table 5.2. Teacher educators' dominant conceptions including operational definitions, illustrative excerpts and frequency. 69

Table 5.3. Criteria categories held by teacher educators with operational definitions, illustrative excerpts and frequency. 71

Table 5.4. Teacher educators' dominant criteria with their operational definitions, illustrative excerpts and frequency 90

Table 5.5. Categories regarding inclination of teacher educators' line placements for Task 6 and Task7, illustrative excerpts and frequency of each category. 93

Table 5.6. Categories for inclination of teacher educators' line placements for Task6-7... 96

Table 5.7. Teacher educators' slope conceptualizations along with their dominant conceptions of informal line of best fit. 107

LIST OF FIGURES

Figure 5. 1. TE1's line placement on Task 5. 77

Figure 5. 2. TE7's line placement for Task2. 82

Figure 5. 3. TE4's line placement for Task 1. 87

Figure 5. 4. TE4's line placement for Task6. 88

Figure 5. 5. TE6's line placement for Task6. 98

Figure 5. 6. TE4's line placement for Task 6 102

Figure A.1. Golf ball bounce height and drop height..... 142

Figure A.2. Golf ball bounce height and drop height..... 143

Figure A.3. Average ticket cost and total attendance..... 144

Figure A.4. Average ticket cost and total attendance..... 145

Figure A.5. Students' achievement motivation and GPA scores..... 146

Figure A.6. Students' height and shoe size..... 147

Figure A.7. Students' height and birth month. 148

Figure A.8. Angelo's line placement for Task1. 150

Figure A.9. Barbara's line placement for Task1. 150

Figure A.10. Yearly movie attendance and average ticket price. 151

Figure A.11. Yearly movie attendance and average ticket price. 152

Figure A.12. Elementary students' height and shoe size. 152

Figure A.13. Elementary students' height and shoe size. 153

Figure A.14. Elementary students' height and shoe size. 154

LIST OF ACRONYMS / ABBREVIATIONS

MTE	Mathematics Teacher Educator
SKT	Statistical Knowledge for Teaching
TE	Teacher Educator

1. INTRODUCTION

Statistical information ranging from economy to education, from movies to sports or from food to medicine is necessary for statistical reasoning and sense making in personal decisions and in workplace and, using and interpreting data to meet responsibilities as citizens (Franklin, Kader, Mewborn, Moreno, Peck, Perry & Scheaffer, 2005). Therefore, new curriculum standards published in countries such as Australia, Brazil, Costa Rica, South Africa, Spain, the United Arab Emirates and the United States of America advocate the importance of statistical education and include statistics in different educational level instructional programmes (Burril & Biehler, 2011). Among the topics in statistics, statistical association is one of the fundamental big ideas that many researchers and statistics educators recommend to be included in elementary, secondary and tertiary education (Garfield & Ben Zvi, 2004; Burrill & Biehler, 2011). However, research has shown that due to the emphasis on procedural knowledge of concepts in most statistical classrooms, students may not reach understanding of basic principles for statistical association (Garfield&Ben Zvi, 2004; Batanero&Diaz, 2010). Hence, there is a trend in research to introduce inference informally to students as a starting point for the development of formal procedures (Statistics Education Research Journal, November, 2008). In this regard, some national curriculum standards (Australia: Australian Curriculum, Assessment and Reporting Authority, 2012; England: Qualifications and Curriculum Authority, 2007; U.S.A: National Governors Association Center for Best Practices & Council of Chief School Officers, 2010) include informal line of best fit as an introduction to statistical association. Casey (2015a) refers to the informal line of best fit as "fitting a line to data displayed in a scatter plot by eye, without using calculations, or technology to place the line" (p.2). For example, students in grade eight in the United States should be able to model relationship between two quantitative variables having a linear association according to the relevant objective: "informally fit a straight line, and informally assess the model fit by judging the closeness of the data points to the line" (CCSS, 2010, p.56).

There are few research studies examining students' and teachers' conceptualizations of informal line of best fit (Casey, 2015a,b; Sorto, White & Lesser, 2011). These studies mainly give some information about participants' common conceptions or misconceptions of the line of best fit; their predominant criteria for placement of informal line of best fit; accuracy of their fitted lines compared to the line of best fit obtained with the statistical procedures (e.g. least squares regression method or median-median line); and, their interpretation based on the association between quantitative variables displayed on scatterplot. Particularly, Casey (2015a) study done with students regarding informal approach for linear regression pointed to the great importance of the informal line of best fit. She suggested that working on the informal line of best fit prior to the formal introduction of the topic might provide teachers some ideas about their students' prior knowledge for the line of best fit and plan their instruction accordingly. Similarly, Casey (2015b) study conducted with inservice teachers pointed to important implications for their subject matter knowledge for teaching informal line of best fit. Aligned with (Batanero&Diaz, 2010), results suggested that teachers' beliefs, pedagogical content knowledge and instructional decisions depends on their subject matter knowledge. Accordingly, understanding and improving teachers' knowledge of the line of best fit is a necessary condition to enhance their students' related conceptualizations at school (Casey, 2015b). Therefore, in order to put another layer on this research area, building on the results of these aforementioned studies I investigated mathematics teacher educators' conceptualization of informal line of best fit. Such research will be important because knowing the expert approach to the task of informally fitting a line can be informative to improve the teaching of informal line of best fit and formal regression to novice learners (Casey, 2015b). In addition, the introduction of statistical association through line of best fit can be thought as an extension of mathematical lines and their properties, including slope (Casey, 2015b; Casey & Nagle, 2016). Particularly, previous research results raised questions regarding whether individuals' slope conceptualizations have an impact on their conceptualization of informal line of best fit (Casey & Nagle, 2016). Therefore, in this study, I also investigated teacher educators' meaning of slope and its effect on their interpretations for informal line of best fit. In this respect, in this study, I examined the following research questions:

(i) What subject matter knowledge for teaching informal line of best fit do teacher educators have?

- What conceptualization for informal line of best fit do teacher educators have?
- What criteria for placement of informal line of best fit do teacher educators use?

(ii) What slope conceptualizations do teacher educators use for conceptualization of informal line of best fit?

2. REVIEW OF LITERATURE

In this section first I focus on the need to study the concept of informal line of best fit and then mention how researchers explicated the concept. Following, I mentioned the importance of investigating mathematics teacher educators' subject matter knowledge for teaching. Then, I share previous research on students' and teachers' conceptualization of the informal line of best fit. In the final part, I review the effect of slope conceptualization on conceptualization of informal line of best fit.

2.1. The Need for Studying Informal Line of Best Fit

Line of best fit is a model to represent association between two quantitative variables if trend of points in a scatterplot of paired numerical data is similar to a linear shape (Franklin *et al.*, 2005). Since regression analysis refers to the study of association between variables (Franklin *et al.*, 2005; Moritz, 2004), I explain "association" firstly and then "the line of best fit (i.e, regression)" pointing also to the misconceptions students might have. Then, I move to the informal line of best fit.

The concept of association, one of the fundamental ideas of statistics education (Garfield & Ben- Zvi, 2004; Burrill & Biehler, 2011), is considered as an extension of the deterministic concept of mathematical function with an appreciation of the role of variation (Engel&Sedlmeier, 2011; Batanero, Estepa, Godino & Green, 1996). In general, statistical association as a main component in data analysis and probability standards for school mathematics (NCTM, 2000) refers to the correspondence of variation of two statistical variables which are i) both categorical represented in two way frequency tables, ii) one categorical and one interval formulated as the comparison of group; and, iii) both

numerical, also known as covariation, represented in scatterplots using a Cartesian coordinate system (Moritz, 2004). In this regard, the association between two variables use information of one variable to understand, explain and predict values of the other variables (Garfield & Ben- Zvi, 2004). In addition, Burrill and Biehler (2011) stated that "if an association is identified between quantitative variables, regression methods can be used to fit mathematical functions to predict one variable as a function of other." (p.63).

Casey (2014) stated that "linear regression is students' first experience with the fundamental concepts of statistical association and dependence in many cases" (p.1). When there is a linear relationship between two variables, an equation can be computed to provide a mathematical description (Gravetter & Wallnau, 2013) For example, when students apply for admission to a college, most college officials will request personal information documents including Scholastic Achievement Test (SAT) scores. Actually, officials want this information since it has been determined that there is a relationship between SAT scores and college grade point averages (GPA). Particularly, if students have high scores in SAT exam, they tend to be successful in college or if students experience difficulty in SAT, then they are expected to have low grade point average (GPA). Gravetter and Walnau state that "the general statistical process of finding and using a prediction equation for the line that best describes the relationship between two quantitative variables is known as regression, and resulting straight line is called the regression line" (p.558-559).

So, the purpose of regression is to find the line of best fit for given set of data to explain and predict relationship between numerical variables (Garfield & Ben- Zvi, 2004). To reach this goal, students should initially understand what the meaning of "line of best fit" is (Gravetter & Walnau, 2013). Although we can draw lots of different straight lines which seem to be passing through the center of data, it is crucial to find the specific line that gives the best fit to the actual data points (Gravetter & Walnau, 2013). For this purpose, residuals that are defined as the deviations in the y direction between the points in the scatterplot and the placed line should be determined firstly (Franklin, 2005). After that, making the sum of the squares of the residuals minimum is known as least squares method to find the line of best fit or regression line (Hester, 1995). This method is used for getting best fit line equation with determination of slope and y-intercept operationally. The aim of these procedures is to make more accurate predictions and explanations by using

regression line which is as close as possible to the actual data points (Coladarci, Cobb, Minium & Clarke, 2010).

Although mathematically trained people are able to get and use the formulae for the regression line, the meaning of the formulae and the relationship between variables are not easily understood (Engel & Sedlmeier, 2011). Particularly, in most statistics classroom instruction, since more importance is given to these procedures, students have difficulties in making correct statistical reasoning about fundamental big ideas regarding association (Garfield & Ben Zvi, 2004; Sorto *et al.*, 2011) Common misconceptions about statistical association are deterministic, unidirectional, local and causal conceptions of association (Estepa & Batanero, 1996; Estepa & Cobo, 2001; Engel & Sedlmeier, 2011).

Particularly, students having *deterministic conception of association* misconception believe that a correspondence between two variables means assigning only a single value in the dependent variable for each value of the independent variable. For example, according to the deterministic view in a bivariate data set, students tend to think that "X is the single cause of Y and every value x of X, there exists exactly one corresponding value y of Y" (Engel & Sedlmeier, 2011, p.252). They also believe that when such information is not provided, there is no association between variables (Engel & Sedlmeier, 2011). This kind of correspondence between variables is in fact the property of the function from a mathematical view. However, when studying in statistical setting, it is natural to observe several y values corresponding to the same x value because of the inherent variation (Engel & Sedlmeier, 2011). That is, understanding regression depends on an appreciation of the role of variation in data. So, the students who have the deterministic conception of association are not able to reason on the variation of the data.

Students having *the unidirectional conception of association* perceive the association only when the correlation coefficient sign is positive. Otherwise, they interpret inverse association as independence (Estepa&Batanero, 1996; Engel&Sedlmeier, 2011). That is, they think that when the correlation coefficient sign is negative, there is no association.

Students *with local conception of association* use only part of the data to make judgment about the given problem. That is, if the partial information serves to confirm a given type of association, they generalise it to the complete dataset. Particularly, when bivariate quantitative data is represented on the scatterplot, students tend to focus on individual data points (case-oriented view) rather than entire set of points in the data holistically (aggregate view) with characteristics which are not visible in any of the individual cases (Bakker, 2004). For example, some students focus on specific points as first and last, highest and lowest or any subset of the points in the scatterplot of the data to interpret association between variables (Casey, 2014).

Students having the *causation conception of association* believe that there is association between the variables if there is a causal relationship between them. This belief may be the result of instructions that the students receive in mathematics, science and other areas in which all the phenomena are given with a causal explanation (Estepa&Cobo, 2001). Therefore, students reflect their ideas from other areas to the study of association and interpret a strong association between two variables as causal dependence (Estepa&Cobo, 2001).

Researchers emphasize that understanding statistical association requires both students and teachers to identify the distinction between association and causation such that association describes a general trend and associated variables are not necessarily dependent (Casey, 2008).

In order to recognize and overcome such misconceptions on association that may arise in class before students begin the formal regression, Sorto et al (2011) suggests teachers to bring informal line of best fit for class discussion to gather some knowledge on the following: “what methods do students use when asked to draw the ‘line that best fit the data’? Do these methods resemble the method of least square? If they do, how do we build upon this knowledge to develop deeper understanding? If they do not, how do we redirect and justify the methods of least square?” (p.49). This knowledge may make important contributions for teaching least square regression because without a discussion comparing their tendencies about fitting a line with the standard mathematical methods, students

might have difficulty in understanding and interpreting the meaning of regression (Sorto *et al.*, 2011). In addition, understanding students' conception of the informal line of best fit provides teachers anticipating students' ideas prior to introduction of this topic and planning their instruction accordingly (Casey, 2014). Similarly, information of teachers' conception of informal line of best fit might give implications for their subject matter knowledge for teaching informal line of best fit (Casey, 2015b). It is in this respect that in this study I further investigate how teacher educators conceptualize and characterize the informal line of best fit. This might shed also light on what teacher educators emphasize in terms of what teachers should know regarding the concept of informal line of best.

In fact, there is a trend in research to introduce students inference informally before learning formal procedures (Statistics Education Research Journal, November, 2008). Zieffler, Garfield, Delmas & Reading (2008) defined informal knowledge as "either a type of everyday real world knowledge that students bring to their classes based on out-of-school experiences, or a less formalized knowledge of topics resulting from prior formal instruction". It is important to study and consider the role of informal knowledge as a starting point for the development of formal understanding of a particular topic (Zieffler *et al.*, 2008, Gravemeijer & Doorman, 1999). According to this informal approach in statistics, some national curriculum standards (Australia: Australian Curriculum, Assessment and Reporting Authority, 2012; England: Qualifications and Curriculum Authority, 2007; U.S.A: National Governors Association Center for Best Practices & Council of Chief School Officers, 2010) include informal line of best fit as an introduction to statistical association. In particular, the learning of informal linear regression underpin the students' study of formal linear regression and regression of other function forms as quadratic and exponential (CCSSI, 2010). Therefore, the formal line of best fit which is obtained with operational methods (e.g. least square regression) is suggested to be delayed until intuitive notions about linear regression have first been developed (Bakker & Gravemeijer, 2004) by studying informal line of best fit (Casey, 2015a).

Casey (2015a) states the definition of the informal line of best fit as "fitting a line to data displayed in a scatter plot by eye, without using operational methods, calculator or related technology while placing the line" (p.2). This line represents a model relationship between two quantitative variables having linear association such that the model

relationship of the line of best fit can be assessed informally by judging the closeness of the data points to the line (CCSSI, 2010- see Statistics and Probability Standards in the CCSS-M Grades 8- HS).

After all, informal line of best fit is recommended to teach as an initial step of learning linear regression and statistical association (Casey, 2015a,b; Sorto *et al.*, 2011) and included in some national curricula (Australia: Australian Curriculum, Assessment and Reporting Authority, 2012; England: Qualifications and Curriculum Authority, 2007; U.S.A: National Governors Association Center for Best Practices & Council of Chief School Officers, 2010). However, since statistics is a relatively new subject included in curriculum standards, many teachers have not had a possibility to gather knowledge of underlying concepts and principles for practices of data analysis for teaching (Franklin, 2005). Accordingly, understanding and improving teachers' knowledge of line of best fit is a need to enhance their students' related conceptualizations at school (Casey, 2015b). Therefore, Casey (2015b) suggested to study the concept of informally fitting a line with experts in this area to improve teaching of this topic to new learners. Hence, I studied this topic with teacher educators to learn how they conceptualize the informal line of best fit and how they consider the teaching of informal line of best fit before introducing linear regression to improve prospective teachers' knowledge.

2.2. Importance of Mathematics Teacher Educators' Knowledge

Mathematics teacher educators (MTEs) are thought as the educators who are responsible for the education or development of mathematics teachers. In particular, since the teachers should have subject matter knowledge for teaching, MTEs also need to have more than the knowledge mathematics teachers require to provide students to learn mathematics (Beswick&Goos, 2018). In this respect, along with the content knowledge, teacher educators also should be able to reason about how people learn mathematics to enhance students' and also prospective teachers' learning (Zavlavski, 2009). Therefore, MTEs knowledge is of great importance in recent years with attention given to increase the quality of education and teachers (Beswick&Goos, 2018).

For many teachers, learning mathematics are focused on procedural knowledge, in particular, students are scored in high grades in mathematics for their proficiency in using the procedures (Zopf, 2010; CBMS, 2012). However, using procedural knowledge properly does not always indicate students' understanding and learning of the relevant topic. Therefore, it is necessary to teach prospective teachers mathematical knowledge for teaching focusing on 'conception of what mathematics is and how it can be learned' (p.28) (Superfine & Wagreich, 2009). In addition, there is a need for MTEs' knowledge to support inservice teachers for the development of mathematical knowledge for teaching (Masingila, Olanoff & Kimani, 2018). In particular, developing teachers' mathematical knowledge for teaching is important to provide teachers acquiring necessary knowledge of mathematical concepts, skills and procedures along with how and why they work (Zopf, 2010; Masingila *et al.*, 2017).

Even though researchers pointed to the importance of MTE knowledge for teaching mathematics, there is limited research examining the teacher educators' knowledge of teaching teachers (Tzur, 2001; Zavlavski, 2009; Çetinkaya, Erbaş, Çelikdemir, Koyuncu, Kol & Alacacı, 2017). Particularly, one of the studies conducted by Cetinkaya *et al.* (2015) reflected a discussion with some mathematics teacher educators about the extent of MTE knowledge. Though participants' deliberation, the common idea about the mathematical knowledge for teacher educators was such that they should have the knowledge and skills in order to analyze mathematical concepts given in school mathematics curricula and make connections among these concepts included in mathematical content from primary school level to the 1st year of university level. In addition, some of the participants advocated that having experience in teaching mathematics would make a contribution to their understanding of mathematics and also be an indicator of their mathematical content knowledge. On the other hand, Tzur (2001) examined terrain of MTE knowledge development through self-reflective analysis. According to him, through reflection, teacher educators could reach underlying perspectives of their own teaching practice depending on "1) what it means to teach mathematics, 2) how someone comes to know how to teach mathematics and 3) how someone's activities promote others' learning of mathematics teaching" (p.273) that embody the perspectives of mathematics teachers' teaching practice (Tzur, 2001). Along with reflecting their own teaching, conducting research based on

teaching and learning mathematics with students and teachers could also contribute to their understanding of mathematical knowledge for teaching (Zavlavski, 2009).

Regarding to the statistical knowledge, as a part of mathematics education, some mathematics teachers who are in charge of teaching statistical content mentioned the lack of competence for teaching statistics and the need for more courses for this domain in mathematics teacher education programs (Koponen, Asikainen, Viholainen & Hirvonen, 2016). In particular, several studies indicated that teachers focus more on procedural aspects, rather than developing students' statistical thinking in most statistics classroom instruction (Garfield & Ben Zvi, 2004; Sorto *et al.*, 2011; Henrequies & Oliveira, 2013). This situation showed the importance of fostering teachers and preservice teachers statistical knowledge for teaching. As previously mentioned, since the mathematics teacher educator are responsible for education and development of mathematics teachers, examining MTEs' statistical knowledge for teaching could provide teachers necessary information to understand how and why the statistical procedures work and the underlying meaning of the concepts.

2.3. Studies on Conceptualization of Informal Line of Best Fit

There are few studies investigating the conceptualization of informal line of best fit. In fact, two studies focused on students' conceptions of the informal line of best fit prior to the learning of the formal line of best fit; and, one study focused on inservice teachers' subject matter knowledge to teach informal line of best fit.

The first research study by Sorto et al. (2011) was conducted with eighteen university level students who were enrolled to study Industrial Technology and middle school mathematics teaching. The purpose of this study was to reveal these university students' conceptions and misconceptions related to the linear regression before being introduced to the concept in their statistics classes. Building on (Sorto et al., 2011), Casey (2015a) studied 8th grade (ages 11-14) students' understanding of the informal line of best fit prior to the instruction on the formal line of best fit. Casey (2015b) also investigated nineteen teachers' subject matter knowledge for teaching informal line of best fit. Results from these studies informed the field in four related areas: students' and teachers'

conceptions of informal line of best fit; their criteria and methods used to place informal line of best fit, accuracy of their placement and interpretation of informal line of best fit.

2.3.1. Students' and Teachers' Conceptions of Informal Line of Best Fit

Results from all these studies showed that both students and teachers have similar conceptions (Casey, 2015a; Casey, 2015b; Sorto *et al.*, 2011). Particularly, in the first category of conception, students had the meaning of informal line of best fit as *'representing where you expect the relationship between the variables to be'* (Casey, 2015a, p.9). These students interpreted the line as "representing a general relationship between the two variables of interest" (Casey, 2015a, p.9). Also, five out of nineteen teachers described the meaning of informal line of best fit in the similar way that is categorised as "*model*" in the study with teachers (Casey, 2015b). Thus, these participants believed that the line of best fit is a *model* to "show the general relationship between the variables in the population" (Casey, 2015b, p.17).

According to second conception category, students believed that the informal line of best fit has a meaning of *'shows what the data looks like'.* It means that the line indicates where the data points are located on the particular scatterplot (Casey, 2015a). Similarly, Casey (2015b) showed that seven teachers had similar meaning for line of best fit as "the best representation of all the sample data where the line accounts for the data at hand rather than more generalized relationship" (p.17). Such conceptualization referred to the category of '*representer'* for the informal line of best fit (Casey, 2015b).

In the third category, students thought the line of best fit as '*something you used to get close predictions'.* These students believed that the line of best fit can be used to make close predictions for points that are not in the original data set (Casey, 2015a). Also, in study with teachers, two participants utilized the line of best fit as *predictor* with the same purpose as making predictions for values not included in the given data set (Casey, 2015b).

Final conception of students for line of best fit was given as '*average of the points'.* These students defined the line of best fit as an average in between all of the points (Casey, 2015a; Sorto, 2011). Similarly, the teachers in this category also used the line as a bivariate

perspective of determining middle or *typical* values named as average or median (Casey, 2015b).

Different from the students' predominant conceptions, two of teachers had another meaning category of informal line of best fit called as *signal in noise*. The *signal* in noisy processes can be thought as the stability in variable system or the certainty in situations involving uncertainty (Konold & Pollatsek, 2002). In particular, according to Engel and Sedlmeier (2011), the trend in the data as linear regression represents the explained part of the variation whereas the deviation between model and data represents the unexplained variation. Thus, data can be thought as "a structural component plus residuals, that is, Data= Signal+ Noise" (Engel & Sedlmeier, 2011, p.25). Casey (2015b) interpreted this result as encouraging that some teachers held the meaning of signal because researchers advocate this meaning for teachers to have (Konold & Pollatsek, 2002; Engel & Sedlmeier, 2011).

When we compare the results of given studies based on how the participants conceive the line of best fit, teachers and students had similar conceptions in general although teachers were able to give better explanations for their responses because of their background knowledge and experience (Casey, 2015b). In fact, on a conceptual level, teachers are expected to construct a bridge between "a deterministic view of a function and a statistical perspective that appreciates variation" (p.25) in the regression analysis of bivariate numerical data (Engel & Sedlmeier, 2011). However, in the study with teachers (Casey, 2015b), only two out of nineteen teachers had the signal-noise conception supporting functional explanation with appreciation of role of variation in statistical data. Casey (2015b) stated that "teachers' varying conceptions point to some significant gaps in their knowledge" to make sense of statistical data (Casey, 2015b, p.27). Therefore, studying this concept with teacher educators might offer more insight into the conceptions to broad teachers' statistical perspectives for teaching line of best fit.

2.3.2. Criteria and Methods for Placement of Informal Line of Best Fit

Sorto *et al.* (2011) and Casey (2015a,b) also examined both students' and teachers' criteria and methods for placement of informal line of best fit in their studies. According to

the results of these studies, *through as many points as possible* is the most preferred criterion among students (Casey, 2015a) even though it is not dominant criterion among teacher responses (Casey, 2015b). Since students studied linear functions before and conceived line as going through all of the points, some of them struggled to place informal line of best fit to given data points that are not linear (Casey, 2015a). Therefore, this group of students focused on collinear points in dataset rather than general trend of the data to place the line of best fit (Casey, 2015a). In addition, some of them started their lines from the origin with consideration of "that is where all lines start" or knowledge of context for the particular task (Casey, 2015a). Their emphasis on specific points on the scatterplot such as going through origin or collinear subset of data points might be thought as evidences for case-oriented view of data (Bakker, 2004; Casey, 2015b) and also local conception of association (Estepa & Batanero, 1996; Estepa & Cobo, 2001; Engel & Sedlmeier, 2011).

The second criterion, *equal number of points both sides of the line,* was also predominant among students and teachers (Casey, 2015a,b; Sorto *et al.*, 2011). This criterion might be categorised as a generalization of the median concept which is an average value in univariate setting and refers to an equal number of data points on each side of it (Casey, 2015a; Sorto *et al.*, 2011). For instance, some students placed their best fit lines as horizontal or vertical and so there were also equal number of points on each side of the line (Casey, 2015a). These students also focused on one of the variables given on the scatterplot and placed the line of best fit to represent the average of this variable as working in the univariate setting (Casey, 2015b) Hence, these students were not able to understand the association between two variables that are covarying to some extent in the bivariate setting (Casey, 2015a).

The third category of criteria that is in conformance with the standard procedural method, least squares regression, is '*closest to all of the points'* (Casey, 2015b). While it is one of the most preferred criteria identified in teacher responses (Casey, 2015b), it is also used in students' placement of best fit line commonly (Casey, 2015a; Sorto *et al.*, 2011). These participants place their lines to minimize the distance from data points to line with the aim of getting closest line to all points (Casey, 2015b).

The fourth category is the other most common criterion among teachers and university level students: *sum deviation* (Casey, 2015b; Sorto *et al.*, 2011). According to this criterion, participants tried to equate the sum of the deviations for points above the line to the sum of the deviations for points below the line (Casey, 2015b). While the participants are applying this method, they do not necessarily make measurements, however, they make reference to this criterion in placement and interpretation of line of best fit (Casey, 2015b; Sorto *et al.*, 2011).

Lastly, other methods preferred less frequently by students and teachers were '*reflecting the relationship the variables have based on context knowledge', 'through the first and last points', 'halfway between lowest and highest points', 'starting from the first point then maximizing the number of points it goes through', 'through points which are midpoints of different clusters of points* and *'each point above the line has a symmetric pair below the line'* (Casey, 2015a,b; Sorto *et al.*, 2011).

When the results of these studies relevant to criteria and methods for placement are examined, even if students and teachers use similar criteria in general, their most preferred criterion differs. While students use 'through as many points as possible' or 'equal number of points on both sides' as criteria for placement mostly (Casey, 2015a), more teachers prefer 'closest to all points' or 'sum deviation' criteria for their tasks (Casey, 2015b). These results indicate that while students decide their criteria depending on their previous knowledge as linear functions or average concept in univariate setting, teachers apply the criteria of placement in accordance with standard statistical methods (e.g. least squares method). In this study, I also searched for teacher educators' most preferred criteria and methods for placement of informal line of best fit and explanations of underlying reasons for their choices.

2.3.3. Accuracy of Placements

As a third category of results, Casey (2015a,b) assessed accuracy of participants' line placements in comparison with the least squares regression line for particular dataset. According to the results, there is considerable variability in locations of the lines created by students (Casey, 2015a) in contrast to teachers (Casey, 2015b). The reason of the

inaccurate placements may be the result of students' predominant criteria as 'passing through as many points as possible' and 'having equal number of points both sides of the line' (Casey, 2015a). On the contrary, teachers' line placements on the given tasks are very similar overall and close to the least squares regression line even if they prefer to apply different criteria in placement of their lines and hold distinct conceptions for the line of best fit (Casey, 2015b). The reason of teachers' accurate placements may be due to their predominant criteria for informal line of best fit as 'closest to all points' and 'sum deviation' that have similar rationale with least squares regression method.

In contrast to the interview tasks having strong positive or negative relationship, teachers' placement of informal line of best fit shows variation in locations of lines for the task in which the data displayed no association (Casey, 2015b). One of the reasons why they are unable to place the horizontal line might be their expectation about two variables (elementary school students' height and shoe size) to be positively related based on their contextual knowledge of these variables (Casey, 2015b). According to teachers' work on this task, Casey (2015b) suggest to explore other task conditions with non-associated data including less strongly related variables to reduce confirmation bias. Therefore, in this study, I examined teacher educators' line placements in various conditions including the tasks asking for zero-slope line of best fit that indicates no association between variables.

2.3.4. Interpretation of Informal Line of Best Fit

It is significant to note students' and teachers' interpretation of what lines of best fit show about the relationship between variables on given tasks. Most participants interpret their lines as displaying increasing or decreasing relationship based on which way their lines are inclined (Casey, 2015a,b). In addition, four students and most teachers (14 out of 18) interpret the association between variables as an implication of causation (Casey, 2015a,b), although teachers who have taught this topic previously were more likely to be aware of the distinction between association and causation (Casey, 2015b). This result shows consistency between previous studies emphasizing the difficulty of understanding difference between association and causation that is an indication of causation conception of association (Estepa & Batanero, 1996; Engel & Sedlmeier, 2011).

Another important issue for interpretation of best fit line is to decide when and how to use deterministic (known, algorithmic and more rigid approaches) or flexible (non-deterministic, aware of limitations and context-specific) approaches (Casey, 2015b). In the study with teachers, one of the difficulties participants experienced is to decide when they make use of flexibility and role of context in statistics. For example, teachers used the context to determine specific data points (e.g. going through (0,0) according to the context) that should be on the line of best fit, to deduce a trend or association that is not supported with data set actually (e.g. the belief of positive association between variables given with no association), to discard outlier points in data or to focus on subset of data points while fitting the line (Casey, 2015b). However, since there is a deterministic formula as least squares regression method to find the line of best fit, the general purpose of this process should be to minimize the vertical distances between the data points and line rather than making use of context-specific approach (Casey, 2015b). Conversely, context is important for interpreting slope of the line of best fit in statistics (Casey, 2015b). Although teachers have deterministic conception of the slope as believing that for every increase in x there should be a precise impact on y (Engel and Sedlmeier, 2011), the correct interpretation of slope of the line of best fit should be as "a typical, average or expected change and association between two variables" (Casey, 2015b, p.26).

The interpretations for what the line of best fit shows about the relationship between variables indicate that participants of given studies have some thoughts supporting causation and deterministic conception of association. Therefore, this study might give more information to interpret association better and how to use deterministic or flexible approaches in interpretation of linear regression.

2.4. Slope Conceptualization and Line of Best Fit

Slope is a fundamental mathematical concept that is introduced to students with linear functions and important for development of topics such as quadratic and exponential functions in precalculus, limit and derivative in calculus and line of best fit in statistics (Nagle, Casey, & Moore-Russo, 2017). This concept can be represented in various forms in the mathematics curriculum such as algebraically in formulas and equations, geometrically in graphs, trigonometrically as tangent of an angle and in real world

representations (Stump, 2001). Particularly, Zaslavsky, Sela, & Leron (2002) identified two main perspectives for slope: analytic and visual. In the analytical perspective, slope is a functional property that can be calculated by using derivative, difference- quotient ($\frac{Y2-Y1}{X2-X1}$), or the coefficient of x in the equation form as y=mx+n. On the other hand, in the visual perspective, slope is a property of the graph of a function and is computed with tan α or ratio as 'rise over run' in relation to angle α between the x-axis and the line of function. In addition, Moore-Russo, Conner & Rugg (2011) suggested eleven slope conceptualizations based on their own research and previous work of Stump (1999, 2001) (see Table 1).

Table 2.1. Conceptualizations of slope (Moore-Russo *et al.*, 2010, p.9).

Category	**Slope as...**
Geometric Ratio (G)	Rise over run Vertical displacement over horizontal displacement
Algebraic Ratio (A)	Change in y over change in x Representation of ratio with algebraic expressions, $\frac{Y2-Y1}{X2-X1}$
Physical Property (P)	Property of line often described using expressions like "steepness" ("slant", "pitch", etc.) "how high up" or "it goes up"
Functional Property (F)	Constant rate of change between varibles
Parametric Coefficient (PC)	Coefficient m in equation y= mx+b
Trigonometric Conception (T)	Tangent of a line's angle of inclination Direction component of a vector
Calculus Conception (C)	Limit Derivative Tangent line to a curve at a point
Real world situation (R)	Static, physical situation (e.g. wheelchair ramp) Dynamic, functional situation (e.g. distance vs time)

Table 2.2. Conceptualizations of slope (Moore-Russo *et al.*, 2010, p.9) (cont.).

Category	**Slope as...**
Determining Property (D)	Property that determines parallel, perpendicular lines Property with which a line can be determined , if you are also given a point
Behaviour Indicator (B)	Real number with sign which indicates increasing, decreasing and horizontal trends of a line Real number with magnitude which indicates amount of increase / decrease of a line Real number that, if positive or negative, indicates line must intersect the x-axis
Linear Constant (L)	Property that is unaffected by translation Constant property unique to "straight" figures Constant property independent of representation

Since slope concept is studied in linear functions before introduction of informal line of best fit, some questions emerged regarding whether some slope conceptualizations in previous topics as functional property and real world situation positively affect students' conceptions of best fit line and criteria to place it, while others as linear constant conceptualization potentially create limitation on students' perspectives for informal line of best fit (Casey & Nagle, 2016). Therefore, examining how participants' prior knowledge regarding slope concept influence their ability to reason in the context of informal line of best fit might be effective to understand reasoning behind their conceptions and criteria for placement of the best fit line (Casey & Nagle, 2016; Nagle *et al.*, 2017). Particularly, identifying the relationship between slope conceptualization and conceptualization of the line of best fit might be informative for teaching of both topics to provide more connected and supportive learning processes and eliminate potential cognitive conflicts (Casey & Nagle, 2016).

There are two studies investigating the effect of students' slope conceptualizations on reasoning about informal line of best fit. Casey & Nagle (2016) analyzed the data of seven students from the previous study (Casey, 2015a) once again with the aim of looking for the

relationship between students' use of slope conceptualizations and their conceptualization of line of best fit. Then, Nagle *et al.* (2017) utilized the answers of two students in the previous study (Casey & Nagle, 2016) to describe knowledge of slope students transfer to the placement of informal line of best fit in detail. Results of these studies showed that students' use of slope conceptualization had an impact on their perspectives for informal line of best fit (Casey & Nagle, 2016; Nagle *et al.*, 2017). Particularly, common slope conceptualizations that students utilized in their decisions for placement of informal line of best fit were linear constant, behaviour indicator, real world situation, and functional property (Casey & Nagle, 2016).

Linear constant conceptualization of slope is focused on the property that lines are straight. Therefore, when students encounter with statistical setting where data points are not lined up because of inherent variation, students with this slope conceptualization may experience cognitive conflict to engage with given tasks and then result in inaccurate line placements (Casey & Nagle, 2016). On the other hand, some students who use the linear constant conceptualization in conjunction with behaviour indicator and real world situation conceptualizations were able to find the best placement of the line accurately. (Casey & Nagle, 2016; Nagle *et al.*, 2017).

The *behaviour indicator* conceptualization of slope was the predominant conceptualization in determination of line of best fit since all participating students examined given tasks whether the data points in the scatterplot have increasing or decreasing trend (Casey & Nagle, 2016). When the placements of students were analyzed, responses where the lines placed inaccurately were the result of the using only behaviour conceptualization of slope or using it in conjunction with linear constant conceptualization (Casey & Nagle, 2016). In contrast, when the behaviour indicator conceptualization was applied with *visual* as physical or trigonometric conceptualizations or *analytic* as real world situation or functional property conceptualization perspectives of slope (Zavlavski *et al.*, 2002), students reached accurate line placements (Casey & Nagle, 2016).

It is significant that when students used the *real world situation* conceptualization of slope, the placement of line was always accurate (Casey & Nagle, 2016). In fact, this conceptualization of slope reflects the purpose of the line of best fit as representing the

relationship between the variables (Casey & Nagle, 2016). In addition, many students explained their lines as representing covariation between the real world variables and it reveals that these students use real world situation conceptualization with functional property conceptualization (Casey & Nagle, 2016). Thus, this interpretation of line's slope as "describing the rate of change between variables in context" (p.175) is in conformity with recommendations in statistics education (Casey & Nagle, 2016).

On the other hand, the only study investigating teachers' conceptualizations of slope while reasoning about informal line of best fit (Casey, 2015b) looked for teachers' calculations to find the estimated value of slope of their best fit line (for Task 3, see Appendix A) and their interpretation of meaning of slope in context. Analysis focusing on teachers' interpretation of slope showed that they describe the slope of line of best fit as "precise and constant rate of change" (p.22), similar to interpretation of slope in mathematical setting (Casey, 2015b). Particularly, many teachers had deterministic view (known, algorithmic, and more rigid approach) of slope and believed that the changes in the response variable were caused by changes in the explanatory variable in given task (Casey, 2015b). In addition, some teachers did not mention task context (Casey, 2015b), essential for understanding and describing the slope of best fit line (Franklin, 2005). Particularly, slope of a line of best fit in statistical setting is recommended to be interpreted in context situation as an approximation or average value described as "the typical or expected change one would expect to see in average, but never 'always'" (p.22) (Casey, 2015b).

To conclude, slope of best fit line can be described as expected rate of change between the variables in context situation. Therefore, teachers' emphasis on the coordinated use of real world situation and functional property conceptualizations of slope in their instructions might positively affect the initial understanding of line of best fit and provide more cohesive learning process (Casey and Nagle, 2016).

Casey (2015b) examined the teachers' meaning of slope for their informal line of best fit, however, she did not give sufficient information about teachers' slope conceptualizations that affect conceptualization of informal line of best fit. In this study, I also investigated teacher educators' slope conceptualizations while thinking of the line of

best fit to provide important connections for teaching the slope concept in mathematics and statistics. In addition, I examined how mathematics teacher educators calculate the value of slope of placed informal line of best fit and interpret it in context. This information might indicate the role of slope in interpretation of the linear regression.

In the light of these information mentioned above, previous studies focused on the concept of informal line of best fit examining students' and teachers' conceptions and misconceptions regarding linear regression (e.g., Casey, 2015a,b; Sorto et al., 2011). Casey (2015b) emphasized the need to work with experts to further investigate the concept of informal line of best fit: "It would be informative to know how experts approach to the task of informally fitting a line and to use the findings to improve the teaching of this topic to novice learners" (p.27). Therefore, in this study, I aimed at putting another layer on previous research by investigating mathematics teacher educators' subject matter knowledge regarding their conception and criteria for informal line of best fit, and interpretation of association between variables. In addition, the introduction of statistical association through line of best fit can be thought as an extension of mathematical lines and their properties, including slope (Casey, 2015b; Casey and Nagle, 2016). Therefore, in this study, I will be also focusing on teacher educators' slope conceptualizations and its effect on their interpretations for informal line of best fit. In this respect, in this study, I will be examining the following research questions:

(i) What subject matter knowledge for teaching informal line of best fit do teacher educators have?
    - What conceptualization for informal line of best fit do teacher educators have?
    - What criteria for placement of informal line of best fit do teacher educators use?

(ii) What slope conceptualizations do teacher educators use for conceptualization of informal line of best fit?

3. THEORETICAL FRAMEWORK

3.1. Statistical Knowledge For Teaching

This study mainly draws on Statistical Knowledge for Teaching Framework (SKT) (Groth 2007, 2013) to guide research regarding knowledge for teaching statistics. Since statistics utilizes mathematics, the categories such as subject matter knowledge and pedagogical content knowledge in Mathematical Knowledge for Teaching (MKT) framework (Hill, Ball and Schilling, 2008) are used to structure the elements of SKT (Groth, 2007, 2013). In this study, in accordance with Casey (2015b), subject matter knowledge of teacher educators will be examined. Therefore, only this part of the framework will be explained.

Subject matter knowledge includes three categories: common content knowledge, specialized content knowledge and horizon knowledge. Common content knowledge is related to competencies developed in conventional statistics courses and used in the work of many occupations along with use of statistics in teaching (Groth, 2013). In this study, accurately reading scatterplots, finding linear model representation for given data on tasks and correctly interpreting the line of best fit can be classified as common content knowledge for statistics. Therefore, in this study the Tasks 1-7 given with different strength and direction of association are used to investigate teacher educators' common content knowledge regarding informal line of best fit to answer research questions for teacher educators' conceptualizations of the line of best fit.

As part of subject matter knowledge that is not necessary for professions other than teaching practice, specialized content knowledge relates to teachers' knowledge of how to explain and justify the reasoning behind mathematical ideas (Ball, 2008). Therefore, in this study, teacher educators' explanations and justifications related to conceptualization of the line of best fit and association between variables on the Tasks 1-7 could be an indicator of their specialized content knowledge. In addition, using the slope value of informal line of best fit to interpret the association between variables in given contextual information on 9th

and 10th questions is also used to understand teacher educators' specialized content knowledge regarding informal line of best fit. Particularly, teacher educators' knowledge of using slope conceptualizations to interpret the association between variables might give valuable insight to provide connection and coherence in teaching of linear functions in mathematics and the line of best fit in statistics (Casey, 2015b).

Additionally, specialized content knowledge also relates to statistical issues in teaching contexts as "appraising students' unconventional methods for solving problems, and constructing and evaluating multiple representations for concepts" (Groth, 2007, p.428). So regarding the informal line of best fit, in this study, I think of understanding differences in students' line placements on scatterplot in Task 8, and identifying and evaluating students' criteria for placement of the informal line of best fit as examples of specialized content knowledge for teaching line of best fit. Though, Hill and her colleagues (2008) mention teachers' knowledge of students and content as part of pedagogical content knowledge. They relate such knowledge to teachers' knowledge gained through observing students while teaching the concept for some time. Thus, I also acknowledge that depending on the teaching expertise of teacher educators such knowledge might be related to their knowledge of students and content (Hill *et al.*, 2008).

Table 3.1. Data sources used to investigate teacher educators' subject matter knowledge for teaching informal line of best fit.

Subject Matter Knowledge	**Data sources**	**Content knowledge**
Common Content Knowledge	Tasks 1-7	Accurately reading scatterplots, finding linear model representation for given dataset on tasks and correctly interpreting the line of best fit.
Specialized Content Knowledge	Task 1-7	Explanations and justifications related to participants' conceptualization of the line of best fit and association between variables.
	Task 8	Understanding differences in students' line placements, identifying and evaluating students' criteria for placement of line of best fit.

Table 3.2. Data sources used to investigate teacher educators' subject matter knowledge for teaching informal line of best fit (cont.).

Subject Matter Knowledge	**Data sources**	**Content knowledge**
Specialized Content Knowledge	9th and 10th questions	Using the slope value of informal line of best fit to interpret the association between variables in given contextual information.

Horizon knowledge as the last category of subject matter knowledge for teaching statistics is required to "know statistics beyond the prescribed curriculum" (p. 9) and it can help teachers to lead their classrooms to provide learning basis for later grade levels (Groth, 2013). For example, Groth (2013) stated that knowledge of standard deviation might be considered as horizon knowledge for sixth grade teachers studying mean and absolute deviation at this grade level. Since mathematical formulas for standard deviation and mean absolute deviation have similar structures, mean absolute deviation could be considered as a precursor for the standard deviation as a more complicated measure of spread (Groth, 2013). Similarly, Casey (2015b) argued that dealing with non-associated data could be classified as horizon content knowledge for teachers since the line in this condition is not for describing the present data but for making statistical inferences about association or non-association based on slope. However, since making statistical inferences using the linear regression is included in the schedule of university level courses, this knowledge is examined in specialized content knowledge for teacher educators. In particular, teacher educators' content knowledge related to statistical inference, which can be identified in participants' explanations for Tasks 6 and 7 with non-associated variables might provide valuable insights into teachers' horizon knowledge for teaching informal line of best fit (Casey, 2015b).

3.2. Teachers' Conceptions and Criteria Framework for Informal Line of Best Fit

Casey (2015b) utilized SKT Framework in her study regarding teachers' statistical knowledge to generate conception and criteria categories for informal line of best fit. In

this study, I also used these categories to prepare and analyse the interviews to depict teacher educators' conceptions and criteria for informal line of best fit. In addition, I consider slope conceptualization categories (Moore-Russo *et al.*, 2010) to examine teacher educators' slope conceptualizations they used to place and interpret informal line of best fit. I now elaborate on the frameworks (Casey, 2015b, Moore-Russo *et al.*, 2010).

Conception framework for informal line of best fit has five main categories; namely, model, signal, typical, representer and predictor (Casey, 2015b) (see Table 3.2).

Table 3.3. Conception categories (Casey, 2015b, p.17).

Conception	**Description of line of best fit**
Model	The model to show the general relationship between the variables in the population.
Signal	The meaningful signal for data that eliminates the inherent noise.
Typical	The bivariate equivalent for determining middle or typical values; sometimes named as average, other times as median.
Representer	The best representation of all the sample data, where the line accounts for the data at hand rather than a more general relationship.
Predictor	The line that enables predictions for values not in the data set.

Teachers having model conception describes the line of best fit as a line that "closely approximates the actual true relationship as possible" (Casey, 2015b, p.17). In particular, this conception is used to make estimation for general relationship between variables in the population. *Signal in noise* conception involves thinking of each individual value on the best fit line as an estimate of a specific value such that each value "is viewed as deviating from the actual data by a measurement error and the average of these values is interpreted as a close approximation to the actual data" (Konold & Pollatsek, 2004, p.179). Thus, the line of best fit is described as "the meaningful signal for data that eliminates the inherent

noise" (Casey, 2015b, p.17). The *typical* conception is related to the data's central position in a distribution of values (Konold & Pollatsek, 2004). That is, teachers with this conception believe that the line of best fit is "a line that averages all the data points" (Casey, 2015b, p.17). Teachers having *representer* conception refers to the line of best fit as the line that represents all the dataset at hand rather than any particular data point (Casey, 2015b). In particular to this conception, the best fitting line could be used to make estimations for where the data points are situated (Casey, 2015a). Finally, teachers with the *predictor* conception believe that the line of best fit is to get close predictions for values not in the dataset. In accordance with this conception, the line of best fit is utilized to make estimations for unknown or future values. I examined the data to answer the research question "What conceptualization for teaching informal line of best fit do teacher educators have?" by using this conception framework.

Criteria framework for informal line of best fit has four main categories: equal number, pairs, closest and sum deviation (see Table 3.3*).*

Table 3.4. Criteria categories (Casey, 2015b, p.17).

Criterion	**Description of criteria to place line of best fit**
Equal Number	Equal number of points above and below the line
Pairs	Each point above the line has a symmetric pair below the line so that the distance to the line is the same for each member of the pair; pairing off individual data points to determine line of best fit
Closest	The line that is closest to all points (implies total distance from points to line is minimized)

Table 3.5. Criteria categories (Casey, 2015b, p.17) (cont.).

Criterion	Description of criteria to place line of best fit
Sum Deviation	Deviations for points above the line to sum to the same as the sum of deviations for points below the line (they did not necessarily make measurements, but referenced this as they placed their line).

According to *equal number* criterion, participants place best fit line as having half of data points above and the other half below the line. The participants who refer to the criterion as *pairs* determine their line as making pairs of each individual data points. That is, they think of mirrored pairs that have similar distance from the line. According to the *closest* criterion, the participants place their best fit line as close as possible to all data points. This criterion reflects the idea behind the least squares regression method as minimizing the total distance from points to the line. In the last criterion, *sum deviation*, participants determine their best fit line as approximately equating the sum of deviations for points above the line to the sum of deviations for points below the line. These criteria categories will be utilized to find out the answer for the research question: "What criteria for placement do teacher educators use for teaching informal line of best fit?"

3.3. Slope Conceptualization Framework

In this study, I also used the slope conceptualization categories (Moore- Russo *et al.*, 2010) dwelled on (Stump, 1999, 2001). The concept of slope emerges in various forms (see Table 3.4).

Table 3.6. Conceptualizations of slope (Moore-Russo *et al.*, 2010, p.9).

Category	**Slope as....**
Geometric Ratio (G)	Rise over run Vertical displacement over horizontal displacement
Algebraic Ratio (A)	Change in y over x Representation of ratio with algebraic expressions, $\frac{Y2-Y1}{X2-X1}$
Physical Property (P)	Property of line often described using expressions like "steepness" ("slant", "pitch", etc.) "how high up" or "it goes up"
Functional Property (F)	Constant rate of change between varibles
Parametric Coefficient (PC)	Coefficient m in equation y= mx+b
Trigonometric Conception (T)	Tangent of a line's angle of inclination Direction component of a vector
Calculus Conception (C)	Limit Derivative Tangent line to a curve at a point
Real world situation (R)	Static, physical situation (e.g. wheelchair ramp) Dynamic, functional situation (e.g. distance vs time)
Determining Property (D)	Property that determines parallel, perpendicular lines Property with which a line can be determined , if you are also given a point
Behaviour Indicator (B)	Real number with sign which indicates increasing, decreasing and horizontal trends of a line Real number with magnitude which indicates amount of increase / decrease of a line Real number that, if positive or negative, indicates line must intersect the x-axis
Linear Constant (L)	Property that is unaffected by translation Constant property unique to "straight" figures Constant property independent of representation

The first category of conceptualization of slope is *geometric ratio* determined as the ratio of vertical change to horizontal change between two points on the line. This category can also be described as " $\frac{rise}{run}$" on the graphical representation of the line. The second conceptualization, *algebraic ratio*, has the meaning of slope in an algebraic formula "$\frac{Y2-Y1}{X2-X1}$" as 'the change in y coordinates over the change in x coordinates' of two selected points on the line. Participants who define slope as a *physical property* use expressions such as "slant", "steepness", "incline", "pitch", and "angle". In general, when the participants are considering tilt of a line, it means that they are using physical property of slope (Nagle *et al.*, 2013). *Functional property* of slope refers to 'the rate of change' between two variables having functional relationship. This ratio is similar to slope conceptualizations as geometric and algebraic ratio mentioned above. However, functional property as a rate of change can be represented in several ways as formula, tables, graphs, physical situations and verbal descriptions (Stump, 1999). For the *parametric coefficient* conceptualization, slope can be interpreted with reference to *m* in the line equation as $y = mx+b$. For example, if we assume that an equation y = 3x+5 is given, then the slope of this line will be represented with the number 3 as a coefficient of x. The next category, *trigonometric conception of slope*, is described with the tangent of an angle existing between a straight line and x axis. The participants having *calculus conception* refer to the concepts such as derivative and limit, and tangent line to a curve at a point (Moore- Russo *et al.*, 2010). For example, if there exists a parabola $y= x^2$, then the slope of the function will be y=2x at x (Nagle, C., Moore-Russo, D., Viglietti, J., & Martin, K., 2013). *Real world situations* include representations of slope in two different forms: *physical* situations such as ladder construction or placement and wheelchair ramp specifications etc, and *functional* situations such as distance- time and quantity- cost (Stump, 2000). Teachers can find out evidences of students' understanding of slope from their interpretations for these real world situations (Stump, 2000). *Determining property* of slope involves determining if the lines are situated as parallel (if they have the same slope) or perpendicular. Also, according to this property, the slope value can be used to draw a line if there is a given point (Moore-Russo *et al.*, 2010). For the *behaviour indicator* conceptualization, positive, negative and zero slopes are used as an indicator for behaviour of a line as increasing, decreasing or horizontal (Moore-Russo *et al.*, 2010). Also, the magnitude of slope gives information for amount of increase or decrease of the line (Moore-Russo *et al.*, 2010).

Finally, based on the *linear constant* conception, slope is used in cases of "straight" figures (Russo-Moore *et al.*, 2010). According to this view, slope can be calculated by using any pair of the points on the given line because it will be the same in each case (Moore-Russo *et al.*, 2010). I analyzed collected data by using this slope conceptualization framework to answer the research question: "What slope conceptualizations do teacher educators use for conceptualization of informal line of best fit?"

In sum, in this study, I make use of these frameworks to examine teacher educators' conceptions and criteria for placement of informal line of best fit and their slope conceptualizations while thinking of the informal line of best fit.

4. METHOD

In this section, I give detailed information about the method of this study. In particular, I elaborated on research design, participant information, data collection and data analysis for this study. Then, I explain validity and reliability issues.

4.1. Research Design: Basic Qualitative Research

Based on the basic qualitative research model, researchers mainly aim to identify how people make sense of their meanings and experiences (Merriam, 2009). In basic qualitative research, particularly, data is collected through conducting interviews, making observations or/and document analysis and then analysed to derive patterns or themes to understand and interpret participants' meanings for relevant conditions (Merriam, 2009). It is in this respect that I implemented basic qualitative research design in accordance with the purpose of this study since the focus was on how the participants make sense of the informal line of best fit and how they interpret the association between variables on given tasks and questions.

In particular, the aim of this study is to understand mathematics teacher educators' conceptualizations of the informal line of best fit. Therefore, I implemented task based interviews to derive information for participants' conceptions and placement criteria for the informal line of best fit alongside of their slope conceptualizations for this topic.

4.2. Participants

In this study, I investigated how mathematics teacher educators working in different universities in Turkey conceptualize the informal line of best fit. Therefore, first I explained general qualifications of mathematics teacher educators and then I gave necessary information relevant to the participants of this study.

In general, teacher educators are described as the one who do research for the development of teaching practice, learning and curriculum, train preservice teachers to improve their knowledge and skills for teaching and train prospective teacher educators (graduate students) (Cetinkaya *et al.*, 2015; Association of Teacher Educators (ATE), 2006). Particularly, they are responsible for having appropriate subject matter knowledge for training their students and investigating theoretical and practical problems to improve teaching and learning (Association of Teacher Educators (ATE), 2006).

On the other hand, 'mathematics teacher educator' refers to the teacher educators who work in mathematics education department of a university and/or teach courses on mathematics education (Cetinkaya *et al.*, 2015). In particular, since mathematics teachers are responsible for teaching statistical content, so mathematics teacher educators are also in charge of their statistical knowledge for teaching. However, it is important to underline that teacher educators do not necessarily have the knowledge of all of the undergraduate level mathematical/ statistical content for teaching (Cetinkaya *et al.*, 2015; Jaworski, 2008). Besides this, having the teaching experience for particular courses could be evaluated as an indicator for their subject matter knowledge (Cetinkaya *et al.*, 2015).

In this respect, I used purposeful sampling method (Patton, 2015) while chosing the participants in this study. Particularly, I searched for mathematics teacher educators who experienced in teaching undergraduate and/or graduate level statistics and/or quantitative research methods courses that include the line of best fit in their schedule. In this process, I examined all universities in Turkey with mathematics education department. Then, I looked at mathematics teacher educators profiles if they qualified for the selection criteria. Among the teacher educators who shared their profiles I determined 15 mathematics teacher educators through examining their background information focusing on the courses they taught before. Among them, I conducted task-based interviews with 11 mathematics teacher educators who were volunteers to attend to this research study.

Particularly, the participants of this study was 11 mathematics teacher educators, 7 males and 4 females, who are experienced for teaching undergraduate/ graduate level courses relevant to statistics and/or quantitative research at least once. Additionally, five

participants (TE1, TE3, TE5, TE6 and TE11) only taught undergraduate courses, four participants (TE2, TE8, TE9 and TE10) taught both undergraduate and graduate courses, and two participants (TE4 and TE7) taught only graduate courses. Total number of courses that participants taught ranged from 1 to 6.

In addition, I took into account the courses participants attended as a learner in undergraduate/graduate level (See Table 4.1). Particularly, two participants (TE1 and TE6) only attended undergraduate courses, seven participants (TE2, TE3, TE5, TE7, TE8, TE9 and TE11) attended both undergraduate and graduate courses, and two participants (TE4 and TE10) attended only graduate courses as a student. Total number of courses that participants took was ranged from 1 to 6.

Below, I shared detailed information for participants' teaching and learning experiences relevant to the statistics and quantitative research courses including the linear regression in their schedule.

Table 4.1. Participant information about their teaching and learning experiences.

Participants	**Level of Courses**	**The courses they taught**	**The courses they attended as a learner**
TE1	Undergraduate	Introduction to Quantitative Methods (7 times)	Statistics I
		Statistics (3 times)	
		Inferential Statistics (1 times)	
		Teaching Probability and Statistics (2 times)	
	Graduate	-	-
TE2	Undergradute	Research Methods in Maths and Science Education (2 times)	Research Methods in Science and Mathematics Education

Table 4.1. Participant information about their teaching and learning experiences (cont.).

Participants	**Level of Courses**	**The courses they taught**	**The courses they attended as a learner**
		Applied Research in Mathematics and Science Educations (3 times)	Statistics
	Graduate	Advanced Quantitative Methods (5 times)	Experimental Design in Education I
			Experimental Design in Education II
		Applied Advanced Statistics (1 times)	Advanced Behavioral Measurement
TE3	Undergraduate	Teaching Probability and Statistics (1 times)	Introduction to the Probability and Statistics
	Graduate	-	ANOVA
			Mathematical Statistics I
			Mathematical Statistics II
			Statistics For Researchers
TE4	Undergraduate	-	-
	Graduate	Statistics (2 times)	Statistics I
		Advanced Educational Statistics (2 times)	Statistics II
TE5	Undergraduate	Statistics I (3 times)	Statistics
			Statistics for Education
	Graduate	-	Advanced Statistics
TE6	Undergraduate	Introduction to Statistics (1 times)	Probability and Statistics

Table 4.1. Participant information about their teaching and learning experiences (cont.).

Participants	**Level of Courses**	**The courses they taught**	**The courses they attended as a learner**
	Graduate	-	-
TE7	Undergraduate	-	Inferential Statistics
			Statistics
	Graduate	Quantitative Analysis	Statistics
			Quantitative Analysis
			Advanced Quantitative Analysis
TE8	Undergraduate	Probability and Statistics	Probability and Statistics
	Graduate	Advanced Statistical Analysis	Advanced Statistical Analysis
			Statistical Methods in Quantitative Research
TE9	Undergraduate	Statistics I (3 times)	Statistics
		Statistics II (3 times)	
	Graduate	Statistics (2 times)	Statistics
			Advanced Statistics
			ANOVA
			Multiple Regression
			Multivariate Statistics
TE10	Undergraduate	Statistics II	-
	Graduate	SPSS Data Analysis	Quantitative Analysis
TE11	Undergraduate	Probability and Statistics (11 times)	Probability and Statistics
	Graduate	-	Advanced Statistical Analysis
			SPSS Applications

As Table 4.1 demonstrated, participants of this study are selected with attention to the criterion of being experienced in teaching and learning of the courses relevant to statistical association and the line of best fit at least once. This suggests that each teacher educator participated in this study have sufficient knowledge to be able to make interpretation for the tasks and questions included in interview process.

4.3. Data Collection

In this study, task-based interview method (Goldin, 2000) is used to collect data. Task based interviews allow researchers to interpret the participants' thinking processes regarding their knowledge (Goldin, 2000). The interviews were in structured format that includes predetermined tasks and questions given in the interview protocol (Appendix A). However, follow-up questions were also used during the interviews to understand participants' ideas more deeply. The interviews were conducted in the spring semester of 2017-2018 academic year. Each interview was approximately one- hour long for each participant. Participants worked on tasks with scatterplot representation to determine the best fit line for given dataset by means of a music wire. Music wire is chosen as a material for this process because "it is not breakable and is thin (thickness of approximately 0.05 inches) to minimize the issue of obscuring the data points when placed on the scatterplot." (Casey, 2015a, p.6). Also, data sources were transcripts of the video recordings and written artifacts documenting each participant's work on given tasks. Below, I explain in detail the interview protocol and the rationale behind the use of the tasks in the study.

4.3.1. Interview Protocol

The interview protocol in this study consists of eight task-based questions including representation of contextual data with eight ordered pairs on scatterplot to elicit teacher educators' conceptualizations related to the line of best fit. I adopted and modified some of the tasks and questions from Casey (2015a,b) and Sorto et al. (2011) and constructed some others based on the results of these studies. In particular, I adopted the five tasks (Task1-2-3-4-6) and 10th question in this study (Appendix A) from Casey (2015b) and I constructed Task 5, Task 7, Task 8 and 9th question based on the recommendations from previous

studies. Also, the data points on the scatterplot for given tasks were situated as suitable for applying linear models.

In Task 1 and Task 2, the contextual dataset about the height a golf ball bounces when it is dropped from different heights is provided. The variables of these tasks have strongly positive association with r_1=0.96 and r_2= 0.97. The y-intercept of least squares regression line for these tasks is so close to the origin that may give rise to an idea of placing the best fit line passing through the origin (Casey, 2015b). Also, multiple sets of points in data of Task 1 are positioned as close to be collinear. Thus, participants who adopt the criteria of '*going through as many points as possible*' can be noticed from the placement of their line on Task 1. In Task 2, the consecutive points are collinear pairs which have distinct slope values from one to another. Therefore, this position of data may lead participants to prefer the criteria of '*having equal number of points above and below the line*' that can result in a poor line of best fit for this task (Casey, 2015a).

The data in Task 3 display the estimation of the total number of movie tickets sold over an entire year with the estimation of the average price of a ticket for that year based on sales in a sample of cities. On the other hand, Task 4 shows similar data to Task 3 such that it displays data using movie sales from a different set of cities during the same year. The variables of these tasks have negatively strong association with correlation coefficients r_3= - 0.76 and r_4= - 0.92. These two tasks (Task3-4) are different from the previous tasks (Task1-2) regarding to their direction and strength of association. When we take into account of the direction of these tasks, it is not possible to draw a line of best fit which is passing through the origin. Therefore, this situation may lead participants to think about the determination of y-intercept of the line of best fit. For Task3, since there is relatively less strong association between variables (r_3= - 0.76), the participants who apply the criteria as '*going through as many points as possible*' and '*having equal number of points above and below the line*' may get a poor line of best fit when compared with least squares regression line (Casey, 2015a). On the other hand, even though the variables in Task4 have stronger association than Task 3 data, aforementioned methods are also not appropriate to determine the best fit line because the data points are less linearly aligned than in the previous scatterplots (Casey, 2015a). Therefore, since there is a great number of possibility to place line of best fit on Task 4 dataset, positioned as less linearly aligned, participants

need to decide which criteria are more important to determine the line of best fit (Casey, 2015a).

Task 5 represents data about the relationship between students' achievement motivation and their GPA scores. There is a moderate positive association between variables of this task. Also correlation coefficient is calculated as r_5= 0.40, since given sample dataset includes an outlier point (2,100). This task is added in the interview protocol in this study to learn participants' approach and interpretation when they encounter with an outlier point. Also, results of Casey's (2015b) study conducted with teachers showed that teachers have difficulty in finding line of best fit when they are faced with large variability in data. Therefore, I decided to add a task with data that appeared linearly associated but have smaller r value as suggested in the previous research (Casey, 2015b).

The scatterplot in Task 6 displays data regarding the height in inches and shoe size for each of eight students in a teacher's class. In this task, variables are not associated and correlation coefficient is found as r_6 = 0.05. Although the graph indicates that there is no association between variables, one may assume that there is a positive association between the height and shoe size of a person. Participants sometimes make interpretation based on their previous beliefs rather than on presented data (Batanero *et al.,* 1998). Therefore, this task is considered with the aim of identifying the influence of task context and prior beliefs related to association between variables on the placement of the line of best fit (Casey, 2015a). Also, it is included in the interview to get participants' interpretations for statistical inferences in case of nonassociation.

Task 7 displays dataset about the association between elementary students' height in cm and their birth months. As Task 6, there is non-associated variables with correlation coefficient r_7= 0.03 on this task. Even if Task 6 is appropriate to get interpretation for non-associated data, it includes contextual information which based on their previous beliefs may lead participants to think about association between variables positively (Casey, 2015b). Therefore, Task 7 representing non-associated data with unrelated contextual information is added to the interview protocol in this study to get more information about no association.

Table 4.2. Summary of Tasks 1-7.

Task	**Contextual information**	**Direction of association**	**Strength of association**	**Characteristics**
1	Golf ball bounce height versus drop height	Positive	r_1= 0.96	- y-intercept of least square regression line for **Task1-2** is so close to the origin. - Multiple sets of data points on **Task 1**positioned as close to be collinear - The consecutive data points on **Task 2** are collinear points
2			r_2= 0.97	
3	Movie attendance versus average ticket costs	Negative	r_3= -0.76	Relatively less strong association between variables on **Task 3**
4			r_4= -0.92	Less linearly aligned data points on **Task 4**
5	Achievement motivation versus GPA scores	Positive	r_5= 0.40	Dataset includes an outlier point
6	Students' height versus shoe sizes	None	r_6= 0.05	Prior beliefs on task context might lead participants to assume positive association between variables
7	Height versus birth month	None	r_7= 0.03	Nonassociated data given with unrelated contextual information

After the Tasks 1-7, summarized on Table 4.2 above, there are two summary questions to derive detailed information about participants' conceptions and criteria in placement of best fit line:

"Could you tell me what you would say to a student that asked you 'What is the line of best fit?'"

"Could you tell me what you would say to a student to help them draw the line of best fit on a scatterplot?"

Also, there is a question asking for an explanation of what the line shows about the relationship between the variables after Task 1-3-5-6-7 which have different contextual datasets. This question is asked to examine teacher educators' ability to read scatterplots to assess whether a linear association exists between the plotted variables (Casey and Kaplan 2014, cited in Casey, 2015b) and, if there exists an association, to explain the relationship between variables.

Finally, the last task (Task 8) which was prepared by Sorto, White and Lesser (2011) and also adopted by Casey (2015a,b) studies is presented after the summary questions. In this study, I also adopted and extended this task to understand teacher educators' criteria for placing the line of best fit and to elicit their interpretation regarding potential limitations of other criteria applied by students and preservice teachers (Casey & Kaplan (2014) cited in Casey, 2015b). This task includes three parts displaying hypothetical students' line placements (Angelo, Barbara and Jane) for Task1, Task3 and Task 6. Then, participants are asked to choose the better line fits among solutions of these students with different criteria, and explain their reasoning for their choice (Casey, 2015a,b). In the first part of this task (including answers for Task1), Angelo's line goes through the two of the points while Barbara's line does not pass through any points, but is on the closest distance from the all points (it is the least squares regression line) (Casey, 2015b). In the second part for Task 3, the criteria of two students were determined as "equal number of points above and below the line" (Barbara's criteria) and "the line as close as possible to all points" (Angelo's criteria- it is the least squares regression line). In addition, the Angelo's solution is also passing through two of the points in the dataset. Finally, in the last part including Task 6 with non- associated variables, the solutions of three students are given. The placements reflect the students' criteria as "the line as close as possible to all points" (Angelo's criteria- it is the least squares regression line), "going through as many points as possible" (Barbara's criteria- the line is passing through three points) and "reflect prior beliefs on the relationship the variables have regarding context knowledge" (Jane's criteria). The third student's line which is drawn with the intention of reflecting prior beliefs is also passing through two of the data points. The criteria for hypothetical students' line placements are determined in consideration of the common students' criteria for placement in previous studies to be compared with least squares regression line. By

making this comparison, teacher educators might conclude which criterion is more important and appropriate to draw line of best fit.

The remainder of the interview includes 9th and 10th questions focusing on Task1 and Task 3, and asking participants for their interpretations relevant to the slope of the line of best fit on these tasks. In particular, participants are asked to repeat the relationship between variables for Task 1 and Task 3, calculate the slope of their best fit lines for these tasks and interpret the meaning of slope in the given context. In the study of Casey (2015b) with teachers, she asked similar questions only for Task 3. However, she recommended to using task scenarios (such as first two tasks related to drop height versus bounce height) that are more deterministic and have an indication of causal relationship (Casey, 2015b). Therefore, I further included a set of questions based on Task 1. Hence, participants' responses for these questions are evaluated based on the research question examining the participants' slope conceptualizations in parallel with their conceptualizations of the informal line of best fit.

4.4. Data Analysis

For the analysis of data, I used coded analysis (Clement, 2000). Clement (2000) argued that coded analysis

> ...focuses on observations that are assigned to predefined categories by a coder, usually from relatively small segments of a transcript. A transcript is coded when the analyst formulates criteria for recognizing a phenomenon and then lists the places where that phenomenon occurs in the transcript. The conclusions may be at the level of observation patterns alone, or, they can be used as data to support or reject theoretical hypothesis that may have been generated by other means. (p.558)

In this respect, I analyzed the data to support theoretical hypothesis with reference to the three frameworks: conception and criteria categories of line of best fit (Casey, 2015b) and slope conceptualization categories (Moore- Russo et al, 2011).

Since in this study, the main focus was on teacher educators' subject matter knowledge in statistics education, I concentrated on teacher educators' conceptions of the line of best fit, the criteria they used to draw the line of best fit; and, their slope

conceptualization for informal line of best fit. Therefore, for the analysis, I read the transcripts from each participant's interview line-by-line to identify their explanations regarding their conception and criteria of line of best fit and slope conceptualization.

Particularly, reading the interview data line-by-line and using the categories of conception for the line of best fit (Casey, 2015b), first, I examined each participants' conceptions. That is, for each participant, I read the whole interview data for all the task situations. In this regard, participants' statements for task situations during the process of interview were informative to answer research question about their conceptions of the line of best fit. Also, their conceptions of the line of best fit was provided by their responses to the question "Could you tell me what you would say to a student that asked you 'What is the line of best fit?'" As the unit of analysis, I spotted the chunks of data, ranging from a sentence to a paragraph, and coded that chunk with reference to the conception framework (Casey, 2015b). I also searched for evidence of reasoning suggesting possible different conceptions that were not determined previously. After examining relevant data from each participant by determining also different possible conceptions I went back to the whole dataset to cross-compare and contrast the participants' conceptions in total. In this respect, for the determination of participants' dominant conceptions, the frequency of the use of relevant conceptions during the interview process was utilized.That is, after I determined the dominant conceptions each teacher educator depicted, I clustered the data from different teacher educators showing a particular conception. Then, for each task, I read the data from different teacher educators to further conceptualize and determine the conceptions they had depicted within each task. While doing that, I looked for evidences to challenge my conjectures and modify them to cohere with the data. In this sense, I also engaged in constant comparative analysis (Strauss and Corbin, 1990). Below, I shared conception framework on Table 4.3 that is utilized to analyze the data to answer research question "What conceptualization for informal line of best fit do teacher educators have?"

Table 4.3. Conception Framework (Casey, 2015b, p.17).

Conception	Description of line of best fit
Model	The model to show the general relationship between the variables in the population.
Signal	The meaningful signal for data that eliminates the inherent noise.
Typical	The bivariate equivalent for determining middle or typical values; sometimes named as average, other times as median.
Representer	The best representation of all the sample data, where the line accounts for the data at hand rather than a more general relationship.
Predictor	The line that enables predictions for values not in the data set.

Similarly, using the criterion categories for the line of best fit (Casey, 2015b), I analyzed participants' criteria for placement of informal line of best fit. In this process, their criteria utilized on given tasks as well as their explanations for the question 'what he/she would say to a student to help them draw the line of best fit on a scatter plot' was examined to determine their dominant criteria to place their lines. In particular, dominancy was determined based on the frequency of the use of the relevant criteria. In addition, their response to Task 8 as choosing the better line from given student answers and justifications for their responses gave some information about which criterion was more important and appropriate according to their perspectives and their content knowledge to assess students' criteria for the line of best fit. Accordingly, I conducted a similar analysis as in the conceptions. Particularly, I read all data sources from each participant looking for situations regarding their criteria for placement of the line of best fit. In this respect, I adopt and modified the criteria framework (Casey, 2015b) and determined the dominant criterion used by each participant. Then, for coherency, I compared the assumptions and conjectures I came up with by cross-analyzing the data from different teacher educators in all data sources. In this way, going back to the whole dataset, I was able to challenge my conjectures and modify them to cohere with the data. This way, I was also able to determine two more criteria namely reflecting context and ellipse& area approach. Below,

I shared criteria framework on Table 4.4 that is utilized to analyze the data to answer research question "What criteria for placement of informal line of best fit do teacher educators use?"

Table 4.4. Criteria Framework adopted and modified from Casey (2015b, p.17).

Criterion	**Description of criteria to place line of best fit**
Equal Number	Equal number of points above and below the line
Pairs	Each point above the line has a symmetric pair below the line so that the distance to the line is the same for each member of the pair; pairing off individual data points to determine line of best fit
Closest	The line that is closest to all points (implies total distance from points to line is minimized)
Sum Deviation	Deviations for points above the line to sum to the same as the sum of deviations for points below the line (they did not necessarily make measurements, but referenced this as they placed their line).
Ellipse & Area Approach	Constructing a circular shape surrounding all data points, and diving this shape symmetrically to determine line of best fit.
Reflecting Context	Using the expectations and prior beliefs about context variables to determine line of best fit.

For research question regarding slope conceptualization, I examined participants' slope conceptualizations from the 9th and 10th questions in the interview asking for a calculation to estimate the slope of line of best fit and its meaning as well as relevant statements throughout the interview. In this process, I again read the transcripts line-by-

line and coded them with reference to slope conceptualization categories created by Moore- Russo et al (2011). I also searched for evidence of reasoning suggesting possible different conceptions of slope that were not determined previously. In this way, I was able to determine dominant slope conceptions each teacher educator utilized. Then, examining relevant data from each participant, I went back to the whole dataset to compare and contrast the participants' slope conceptions in total. While cross-analyzing the data, I looked for evidences to challenge my conjectures and modified them to cohere with the data. In this sense, I also conducted constant comparative analysis (Strauss & Corbin, 1990). After the cross-analysis, I determined that the teacher educators having particular informal line of best fit conceptions depicted similar slope conceptions. Therefore, I went back to the data from teacher educators with particular informal line of best fit conceptions and re-examined the similarities and differences regarding the slope conceptions. Below, I shared slope conceptualization framework on Table 4.5 that is utilized to analyze the data to answer research question "What slope conceptualizations do teacher educators use for conceptualization of informal line of best fit?"

Table 4.5. Slope Conceptualization Framework (Moore-Russo *et al.*, 2010, p.9).

Category	**Slope as....**
Geometric Ratio (G)	Rise over run Vertical displacement over horizontal displacement
Algebraic Ratio (A)	Change in y over x Representation of ratio with algebraic expressions, $\frac{Y2-Y1}{X2-X1}$
Physical Property (P)	Property of line often described using expressions like "steepness" ("slant", "pitch", etc.) "how high up" or "it goes up"
Functional Property (F)	Constant rate of change between varibles
Parametric Coefficient (PC)	Coefficient m in equation y= mx+b
Trigonometric Conception (T)	Tangent of a line's angle of inclination Direction component of a vector
Calculus Conception (C)	Limit Derivative Tangent line to a curve at a point

Table 4.5. Slope Conceptualization Framework (Moore-Russo *et al.*, 2010, p.9) (cont.).

Category	**Slope as....**
Real world situation (R)	Static, physical situation (e.g. wheelchair ramp) Dynamic, functional situation (e.g. distance vs time)
Determining Property (D)	Property that determines parallel, perpendicular lines Property with which a line can be determined , if you are also given a point
Behaviour Indicator (B)	Real number with sign which indicates increasing, decreasing and horizontal trends of a line Real number with magnitude which indicates amount of increase / decrease of a line Real number that, if positive or negative, indicates line must intersect the x-axis
Linear Constant (L)	Property that is unaffected by translation Constant property unique to "straight" figures Constant property independent of representation

4.5. Trustwortiness: Validity and Reliability

To be able to ensure trustwortiness in this study, I examined appropriate strategies for dealing with internal validity (i.e. credibility), reliability (i.e. consistency/ dependability) and external validity (i.e. transferability). In the following paragraphs, I addressed these issues and how they are implemented in this study.

Internal validity (i.e., credibility) deals with investigators' observing or measuring what they think they are measuring (Merriam & Tisdell, 2015). Since in this study I investigated the conceptions and criteria regarding the line of best fit, I adopted and modified some of the tasks and questions from Casey (2015b) and Sorto et al. (2011) and constructed some others based on the results of these studies. In particular, I adopted the five tasks (Task1-2-3-4-6) and 10th question in this study (Appendix A) from Casey (2015b) and I constructed Task 5, Task 7, Task 8 and 9th question based on the recommendations from previous studies. In particular, Task 5 was added in the interview protocol to learn participants' approach and interpretation when they encounter with outlier points in dataset (Casey, 2015b); Task 7 was added to to get more information about non-

associated data given with unrelated context variables; Task 8 was adopted and extended to elicit teacher educators' interpretation regarding hypothetical students' line placements for Task1, Task3 and Task 6; and 9th question was added to get participants' interpretations for slope of the informal line of best fit with use of different task scenarios.

As another way to verify internal validity, Merriam & Tisdell (2015) suggested the use of expert evaluation. Therefore, after I generated the interview protocol, two experts from mathematics education department including one with statistics education background examined the tasks and questions to evaluate if the tasks were appropriate to investigate research questions. Based on their feedback and suggestions, I finalized the interview protocol.

For the internal validity (i.e., credibility) regarding the results, firstly, I asked a mathematics teacher educator to examine the results based on the research questions to evaluate congruency of emerging findings with the raw data, and tentative interpretations (Merriam & Tisdell, 2015). Secondly, I compared and cross-checked the interview data collected from teacher educators with different conceptions and criteria. This also refers to the triangulation using multiple sources of data (Meriam & Tisdell, 2015).

Regarding the reliability, '*whether the results are consistent with the data collected'* (p.251) need to be determined (Merriam & Tisdell, 2015). Lincoln and Guba (1985, cited in Merriam & Tisdell, 2015), conceptualized reliability in qualitative research as "dependability" or "consistency." In particular, Meriam & Tisdell (2015) stated:

> … rather than demanding that outsiders get the same results, a researcher wishes outsiders to concur that, given the data collected, the results make sense—they are consistent and dependable. The question then is not whether findings will be found again but whether the results are consistent with the data collected. (p.251)

Therefore, in this study, I used multiple tasks while collecting the data to obtain consistent and dependable data. In addition, as another strategy for ensuring the reliability, I opened the raw data for the discussions with experts regarding the congruency of emerging findings and tentative interpretations.

Finally, external validity (transferability) “refers to a description of the setting and participants of the study, as well as a detailed description of the findings with adequate evidence presented in the form of quotes from participant interviews and documents” (Merriam & Tisdell, 2015, p.257). In order to enhance the possibility of the results’ transferability to other settings, in this study, I provided rich and thick descriptions of the participants as well as the findings with excerpts and written documentations from the participants.

5. RESULTS

In this section, I shared the findings relevant to the teacher educators' conceptions and criteria for placement of the line of best fit, and their slope conceptualization while working on the informal line of best fit. In this respect, in each category, I firsty demonstrated all the conceptions and criteria held by teacher educators. Then, I focused on dominant conceptions and criteria referred by teacher educators and explained them in detail. Also, further explanation is given to understand teacher educators' perspectives for nonassociated dataset on Task 6 and 7.

5.1. Conceptions

In this part, I aimed to elaborate on teacher educators' conceptions for the line of best fit. In accordance with this purpose, I focused on three data sources: statements from line placement process for different task situations; the summary question asked as "Could you tell me what you would say to a student that asked you "What is the line of best fit?""; and, general interpretations during the interview. In the analysis of these data, I firstly determined conception categories teacher educators referred throughout the interview process. Then, I identified dominant conceptions for each teacher educator according to the frequency of their reference to these conception categories.

In the analysis of teacher educators' statements, I found evidences for five conceptions as *model, signal, typical, representer* and *predictor*. These conception categories are responses to the research question: "What conceptualization for informal line of best fit do teacher educators have?" In particular, *model* conception was depicted by three teacher educators (TE1, TE7 and TE10), *signal* conception was used by three teacher educators (TE2, TE9 and TE10), *typical* conception was referred by three teacher educators (TE4, TE6 and TE10), *representer* conception was depicted by eight teacher educators (TE1, TE3, TE4, TE5, TE6, TE7, TE8 and TE11) and *predictor* conception is used by eight teacher educators (TE1, TE2, TE3, TE5, TE7, TE8, TE9 and TE11). Below,

I constructed Table 5.1 representing five conceptions held by teacher educators, their operational definitions, illustrative excerpts for each category and frequency of participants who referred to them.

Table 5.1. Teacher educators' use of different conceptions including operational definitions, illustrative excerpts and frequency.

Conception	**Definitions**	**Example excerpt**	**f**
Model TE1, TE7, TE10	-the model to show relationship between variables that could be generalized to the population -the model to show general relationship between variables in the population (Casey, 2015b)	TE1-"The line that fits the best…That is to say, if I think that this is linear, to be able to say that there is a linear relationship…*I mean to be able to see if there is a linear relationship or not, we are trying to find the line of best fit."* (1st summary question) TE7-"...What is my purpose for best fitting curve? To find the relationship between two variables....*It could be used to understand how the relationship exists between two variables.*" (Task 5) TE10- "I will have looked at the relationship between these two (variables). *I mean, which line can explain the relationship between these two the best…*" (Task 1)	3
Signal TE2, TE9, TE10	-Thinking of each individual value on best fit line as an estimate of a specific value such that each value "is viewed as deviating from the actual data by a measurement error and the average of these values is interpreted as a close approximation to the actual data (Konold and Pollatsek, 2002) -Stability in variable system (Konold and Pollatsek, 2002) -the meaningful signal for data that eliminates inherent noise (Casey, 2015b)	TE9-"…We also make them calculate the residuals. We let them see this and that *it's different for each person. For instance Ali got 50 from the exam, the estimate value is his getting 60, that is the difference in between. We show them the value and the errors.* We say that *the value which makes the residuals minimum* but we also say that it can be calculated for each person if necessary." (1st summary question) TE2-"In other words, they (students) have difficulty in this: *The notion to extract a linearity from a data that is not linear* is a difficult issue for them to understand. *What we try to do when we explain the line of best fit is to explain that we are looking for an order in the mess of the data and that one of the best way to express this search for an order is line…*" (summary question) TE10-"We are looking for linearity and there is no linearity that is normal and social in all reason. No factor in bilateral relations… *let's assume we eliminate all. We wonder what kind of a relationship exists among these, it's actually an assumption.* These are totally about an assumption, *can I get a relationship among these if I rule out the other variables*?, If I could how would it be? it would be linear, it would be a factor that affects one to one. However, *since here other variables also affect, and since the point we focus on is two variables, we are trying to eliminate other variables and minimize the*	3

Table 5.1. Teacher educators' use of different conceptions including operational definitions, illustrative excerpts and frequency (cont.).

Conception	**Definitions**	**Example excerpt**	**f**
		*errors***.**" (Task 3)	
Typical TE4, TE6, TE10	-the bivariate equivalent for determining middle or typical values (Casey, 2015b) -the line that averages all data points (Casey, 2015b) or all possible line placements	TE4- "It means that *the line that passes through the middle and is placed in most optimally and fairly for all data points*." (summary question) TE6-"In the end, here we are trying to place the line that fits the data set best. Therefore, I think that it should be something like the points that are closest to it or *the average of the points*. That's why I placed it like this." (Task 4) TE10-"*The regression here gives us an average part of the linear lines that are formed by these (data points),* it's necessary to think this way. If we assume that there is a linearity like this, the first line should go this way, I mean I think that there is a linearity like this (10.8.6.1) (as passing through 1st data point), the situation should be like. But our second point again have an error here. This is not linear. *In fact we gather them all together and find a common intersection set.* This is a linear approach, it is our regression. According to my predictions, we cannot make a calculation of these on a computer one by one so we need to make a linear thing as a common component of them all."(Task 1)	3
Representer TE1, TE3, TE4, TE5, TE6, TE7, TE8, TE11	-the best representation of all sample data, where the line accounts for data at hand, rather than more generalized relationship (Casey, 2015b)	TE1-"Now I have some data at hand. If we think that *this data has a linear relationship, or more precisely we are trying to see if it has a linear relationship. To be able to see this I am looking at the scatterplot of the data I have*. I'm trying to find the line that passes the closest to these points."(1st summary question) TE3-"*The line that best represents the trend of the given bivariate data."* (1st summary question) TE4-" It is the line that *best, most optimally, represents whole set of data."* (1st summary question)	8
Predictor TE1, TE2, TE3, TE5, TE7, TE8, TE9, TE11	-the line that enables predictions for values not in the dataset (Casey, 2015b).	TE1-"Why is this useful? If these points are really close enough to the line, then I will find first the equation of the line in my work I mean in the predictions – you can think of it as regressions – which I did by using this line. After that, *I will be able to make predictions for the values which I didn't have when I wrote x and y on their places. Then, it will give an approximate value, an estimated value. However, the closer my points are to the line, the more consistent my predictions will be*." (1st summary question) TE2-"*We are trying to deduce a prediction from the data. When we talk about the data at hand, we can see the results but we are trying to make predictions about the data that don't exist.* In order to do this, we generally assume the data as linear, we	8

Table 5.1. Teacher educators' use of different conceptions including operational definitions, illustrative excerpts and frequency (cont.).

Conception	Definitions	Example excerpt	f
		try to fit in a linear model. We need to draw a line in the linear model that we are trying to settle. We need equations, functions for the operation that we call prediction and regression. We are trying to represent these functions with a line because it is easier to do so. We are looking for the line which fits the data and which produces the least errors. That is the line of best fit…"(1st summary question) TE3- "This x variable – the height where the ball was dropped - means a very good predictor of the y variable. In other words, *by knowing the value of one of the variables we can predict the second one with little error*." (Task 1)	

After that, I also determined dominant conceptions for each teacher educator based on the frequency of their reference to related conception. In particular, three conception categories were emerged as dominant conceptions held by teacher educators: *representer predictor* and *signal.* In particular, five teacher educators (TE1, TE4, TE6, TE7 and TE8) depicted *representer* conception dominantly, four teacher educators referred *predictor* conception dominantly (TE2, TE3, TE5 and TE11) and two teacher educators depicted *signal* conception dominantly (TE9 and TE10). Since the focus of this study is to understand characteristics of conceptions held by teacher educators dominantly, in the following paragraphs; I elaborate on the data regarding dominant conceptions teacher educators depicted.

Representer: Representer conception was the most frequently used conception for this research study. It is obtained from statements of five teacher educators (TE1, TE4, TE6, TE7 and TE8) as dominant. In accordance with the aim of understanding characteristics of representer conception regarding the informal line of best fit, I firstly share example excerpts from statements of these teacher educators who dominantly used representer conception throughout the interview. Then, I summarized commonalities of their interpretations regarding representer conception.

For Task 1, TE6 and TE8 referred to the *representer conception*. In particular, TE6 stated that:

"...I think the *line should represent them (data points) most appropriately*...Let's say *it passes through 2-3 of them (data points), for instance. After that, it's close to the other ones.* That's what I thought."

Similarly, TE8 pointed that:

"...*I mean this can be the line that represents them (data points).* Otherwise a more special... I'm not affected by the context at the moment. If it was something different from the ball, it would be the same situation. What is important here is that this line cannot be vertical or horizontal. It's essential that vertical distances of these points is minimized...*I think that this is the best line that can represent this scatter of data points."*

As excerpts suggested, TE6 and TE8 defined the line of best fit on Task 1 as the line that represents the dataset. In particular, they focused on the data points at hand and placed the best fitting line as close as possible to all data points. In this respect, TE8 pointed that he was not affected from contextual information and general relationship between given variables. In particular, he minimized the vertical distances to find the best line placement that can represent the scatter of data points.

In addition, TE4 and TE7 held *representer conception* while working on Task 3. Particularly, TE4 stated that:

"Here, I wouldn't tell the students or the prospective teachers to draw a piece of wood or to create an area. I would make a brain storming by asking them *what is the line that best represents the dataset at hand and how do we find it.* I hope some people would suggest this method."

As data showed, he explained the reasoning for his line placement criterion focusing on finding the best representation of the dataset at hand. Therefore, he mentioned that he would lead to students as asking 'what is the line that best represents the dataset at hand and how do we find it?' while determining the line of best fit on given scatterplot.

Similarly, TE7 mentioned that:

"What is my purpose? Determining the best fitting curve...Therefore we always look at the distance. But my best fitting line curve, I can draw it also like this (showing with hand). But *it doesn't represent these then. In order for it to represent,* I need to look at the deviation since I cannot say equal distance or minimum distance. Because it will be linear, it will be on this, won't it? It would be on the line. If it is not on the line, if you say it's as close to the line as possible, than it would be the one that passes through the most points. *But then it wouldn't represent these. It is about representation*. Deviation. Isn't the logic of the deviation is to scatter and to spread? What do we do when we calculate the standard deviation? We calculate the distribution and dispersion. We can think of it as the data's minimum dispersion, minimum distribution and minimum dispersion around the line. So what does that mean? We can think of it as best fitting to these, I suppose."

As excerpt indicated, she reasoned on the criteria for line placement with aim of representing given dataset. In particular, she stated that if the data points were not lined up, then the best fit line would be placed with minimum deviation from given data points in order to represent the dataset. Hence, she summarized her reasoning on line placement process as "it is about representation" based on the *representer conception.*

For Task 5, TE7 and TE8 also referred to the representer conception. In particular, TE7 commented on Task 5:

"After I have drawn this, I think it won't matter if I've drawn it negative or positive. *It does not represent the data... This (2,100) is too high in the deviation*, I will slide up while trying to balance this. *How can I explain this? It won't represent.* In order to equalize the deviation over there, maybe it would be something like this, I'm just making up. I made up too much tough. It would slide up."

Data indicated that, she aimed to represent given dataset as a whole rather than focusing on any particular data point. However, there was a point (2,100) on this task that was far away from general trend of dataset. Therefore, she thought that the best fitting line would not represent whole dataset since this data point (2, 100) was high in deviation. Hence, she also examined Task 5 referring to the *representer* conception.

Similarly, TE8 interpreted Task 5:

"...Is there a line that represents these (data points) like the regression line y = ax + b and that minimizes sum of the squares of errors? Even though the sum of the squares of errors is very high, it has a minimum value, there is a line which provides minimum."

As data suggested, he interpreted that the association between variables was lower than previous tasks. However, he stated that although the sum of squares of errors was too high, there would be a line that minimizes these values. In particular, he described the regression line that provides the minimum errors as a line that represents data points.

Additionally, for Task 7 including nonassociated dataset, TE1 and TE6 used representer conception. In particular, TE1 explained that:

"…Therefore I went back to my points, *I look at the graph I mean. When I look at the graph, when I place it like this*, five points were quite close to the line. So, I *think this is the line that fits best."*

As the data suggested, TE1 determined the best fit line focusing on the graphical representation of data points at hand, dealing with whole dataset. This suggested that she made interpretations on Task 7 revealing the representer conception.

Particularly, TE6 stated that:

"*I intended to draw a line that will represent the data set in the best way* by looking at where the data aggregates and densifies."

Data pointed that, since dataset given as nonassociated, she examined the dataset focusing on where the data aggregates. By doing this, she intended to place a line that will represent the dataset in the best way.

Lastly, teacher educators having the *representer conception* explained 1st summary question as "Could you tell me what you would say to a student that asked you "What is the line of best fit?"" as follows:

TE1 stated that:

"Now *I have some data at hand.* If we think that this data has a linear relationship, or more precisely *we are trying to see if it has a linear relationship. To be able to see this I am looking at the scatterplot of the data I have.* I'm trying to find the line that passes the closest to these points."

As data suggested, she explained that she is focusing on the scatterplot of data at hand to show the linear relationship between variables. Therefore, this explanation pointed to representer conception of the line of best fit.

Also, TE4 mentioned:

"It is *the line that best, most optimally, represents whole set of data."*

Data pointed that he clearly defined the line of best fit as the best representation of whole dataset. Hence, he referred to the representer conception while explaining the summary question.

In addition, TE6 stated that:

"I would ask them to draw *the line that would represent the data set in a best way* by giving examples like you have given here. I usually proceed with this logic."

As excerpt indicated, she thought to direct her students to find the line that would represent the dataset in a best way. Hence, she also used the representer conception to explain the meaning of the line of best fit to her students.

Similarly, TE7 mentioned:

"*The word represent…*I shall first think about the keywords and then I will use it in a sentence. *Represent means representing the data.* Of course you can look at the context to see if it fits but *it's important that it represents the data. While representing, deviation…*I would say that is the balance of the distance of the points that are above the line and below the line. *So I can represent it.* Trying to balance a deviation that will stay close to the points."

As data suggested, her reference to the keyword "represent" indicated that she aimed to represent the data as a whole while placing the best fit line. Therefore, she used the sum deviation criteria as balancing the distances of the points above and below the line in order to represent the data points at hand.

In sum, participants having *representer conception* as dominant referred to this conception for different task scenarios and 1st summary question. Particularly, they described their line placement process using the keyword "represent"in their statements. In this respect, they generally defined the line of best fit as 'the best representation of given dataset'. Particularly, instead of dealing with the general relationship between context variables, these participants gave importance to data points at hand to determine the direction of association and the best fit line. In this respect, they tried to minimize residuals and control deviations around the line to reach the best representation of given dataset as a whole.

Predictor: Predictor conception is obtained from four teacher educators' statements (TE2, TE3, TE5, TE11) dominantly. In order to describe characteristics of the predictor conception of the line of best fit, I firstly present illustrative excerpts from statements of

teacher educators who dominantly referred to this conception during the interview. Then, I elicit the meaning of predictor conception in the light of teacher educators' perspectives.

For Task 1, TE3 and TE5 referred to the predictor conception. In particular, TE3 stated that:

"There is a very strong positive relationship, if I were to predict, I would say there is a correlation close to r=0.90...*This x variable – the height where the ball was dropped - means a very good predictor of the y variable.* In other words, *by knowing the value of one of the variables we can predict the second one with little error."*

As excerpt suggested, he interpreted that there is strong correlation between x and y variables on Task 1. Then, he continued that x variable – drop height of golf ball – could be thought as a predictor for y variable - bounce height of golf ball. In particular, he further explained that he could make predictions for one of the variables with minimum error by using the value of other variable. Therefore, he reflected *predictor* conception of best fit line on Task 1.

Similarly, TE5 stated that:

"This line shows us a relationship that provides making predictions for bounce height of golf ball with using drop height"

As data suggested, TE5 asserted that one could make predictions for bounce height of the golf ball based on its association with drop height. Hence, TE5 addressed the predictor conception as referring to the utility of best fit line that shows a relationship between variables.

In addition, comparing Task 1 and Task 2, TE2 referred to the the *predictor conception* to determine the values out of given data points using the line of best fit:

"…if we will come up with a regression from these (data points), *to be able to make predictions for higher values,* I should see how the trend goes on ..*things out of dataset*… So, I need to examine how their scatter is. That means I need to look at the distribution of each y values corresponding to each x."

As excerpt suggested, he mentioned making predictions for future values of variables with use of regression. For this purpose, he suggested examining scatter of data points for higher values to understand the trend of dataset. This indicates his reasoning regarding *predictor* conception for future values of variables.

In particular, for Task 3, TE11 used the predictor conception as follows:

"...let's say the data are in very far points, then I can't interpret the relationship in between very well. Therefore, *I decide it to be the closest…I do that to make better predictions."*

As the data pointed, she determined the line of best fit as close as possible to the data points. In this respect, she aimed to make better predictions with use of this line.

For Task 5, TE2 and TE3 depicted predictor conception of the line of best fit. In particular, TE2 stated that:

"*…this relationship's being weak shows that there is high margin of error in my predictions,* in the line of best fit that I came up with and *in the predictions I made with this line. In the other ones (tasks), I will have a lower margin of error in the predictions I make.* But *the errors in my predictions here will be very high…* I mean, for *some of them (points) I will make correct predictions, for some of them I will predict wrong and I will find higher error, and for some of them I will predict the value much lower than what actually it is.* Therefore, I already know the data points for such a weak relationship, but *if I will regress it to others, if I will make predictions, it won't serve as a strong prediction for me,* this will be its difference from other tasks."

Data pointed to important issues: First, TE2 compared Task5 to the other tasks by stating that since the data points in this graph were spread out, the relationship between the variables would be weak compared to the other graphs shown in other tasks (e.g., Task1

and 2). He further pointed that such dispersion would not only let him determine the strength of the relationship but also the variability in the predicted data.That is, the predicted values would have more deviation from the actual data points and this would lead to more errors in predictions. Therefore, for him, the more spread out the data points would mean the weaker the relationship; and, the weaker the relationship would be the more variability within the prediction.

Similarly, TE3 interpreted that:

"Of course, it doesn't have correlation coefficient as high as the previous ones, it has lower. If we were to do some more advanced work here, if we were to make statistics in the undergraduate level, we would do hypothesis testing, we would do a test to see whether the β_1 equals to zero or not. What I mean by β_1 is slope here, $y' = \beta_0 + \beta_1 x$. That is to say, in this regression equation, I mean, we find β_0 and β_1 values according to the line of best fit. What would we do at the end is to look whether this β_0 and β_1 are different from 0 or not. What would it mean if it were different from zero or if it equaled to zero? Are we aiming to find the motivation or GPA of the students here? We could take them all and find the average right?...I mean we would predict the average for each student.*What we are doing with the benefit of the line of best fit is trying to predict the GPA from the line we draw and from the y values of equation we will create according to it, instead of predicting it from the average of the students' GPA scores. Therefore, these y values that we will obtain here will give better results than the y values that resulted in taking average directly."*

In given excerpt, he stated that the association between variables was not strong as if previous tasks. Therefore, he mentioned β_1 (slope value of regression equation) hypothesis test that is used to understand whether there exists an association or not and decide type of analysis from among: 1) taking average of relevant variable (if β_1 was equal to zero) or 2) using line of best fit and regression equation (if β_1 is different from zero) to make better predictions. According to TE3, if the association between variables for bivariate data is not meaningful, then there is no need to apply regression methods. In this case, TE3 suggested to taking an average of values for the variable that we want to make predictions for and interpret the result for univarite data. On the other hand, he explained the purpose of using line of best fit to make better predictions for students' GPA scores than predictions based on taking average of GPA scores for given dataset. Then, he continued that since there is a positive association between GPA scores and achievement

motivation, using the line of best fit would provide better predictions. Hence he preferred to use regression model for bivariate dataset based on the *predictor conception* in this task.

Lastly, TE2, TE5 and TE11 commented on 1st summary question based on the predictor conception. Particularly, TE2 stated that:

"*We are trying to deduce a prediction from the data.* When we talk about the data at hand, we can see the results but *we are trying to make predictions about the data that don't exist.* In order to do this, we generally assume the data as linear, we try to fit in a linear model. We need to draw a line in the linear model that we are trying to settle. *We need equations, functions for the operation that we call prediction and regression.* We are trying to represent these functions with a line because it is easier to do so. We are looking for the line which fits the data and which produces the least errors. That is the line of best fit…"

As the excerpt shows, he thought that he could reach the results for existing data points on scatterplot, however, he pointed to a need for a linear model and an equation to make predictions for values that are not included in dataset. Particularly, he aimed to get a line that fits the data with minimum error from actual data values. Thereby, this line would be the line of best fit to make predictions for the data that do not exist.

Similarly, TE5 stated that:

"*This line can be used to make predictions*. For instance, if the independent variable is this, what would be the expected dependent variable value? *This line will make it is easier for us to predict this."*

Given excerpt depicted that he declared the purpose of using the best fit line as to make predictions. That is, for him, the line of best fit allowed him to predict the expected value. However, his use of the words *"...dependent variable*" suggests that he might have been adopting the causation view rather than the association view.

In addition, TE11 mentioned predictor conception for the same question:

"I would explain to him like this: *you are trying to predict something*. For instance, I would give such a data set as an example: a child studied for 3 hours and got 10 from the exam, he studied 4 hours and got 7. Then I would ask whether there is a relationship between these or not…Let's give an example. Let's draw this, and now let's draw it like this, let's draw it like this (she is showing a positive-negative sloped line with her hand), *who will make the best prediction?* If the child gets closer to the most correct one, we can give an example and a data or he can say for each one that, the child studied 8 hours and got 13, will it give that result? Try it and see how close it gets. Such as, he wrote and found 13.1, he wrote the other one and found 13.5. *We want him to find the closest result* by trying on the linear equations that he has found. Because there is a ready data in front of him, he can compare it with that. Therefore an activity can be done like**:** *which one made it easier, so the line which the distance of these points' to each other, to the line is the smallest helped my prediction."*

As data indicated, she explained the use of the best fit line as making predictions for given variables. Then, she clarified it with an example of bivariate data that is assumed to have an association between variables. After that, she hypothesized to lead students to find line placements that fit given dataset. Since she believed that better predictions could be provided with the best fit line, she thought to bring the students to the idea of having closest line to data points that could enable better predictions. Hence, he referred the *predictor conception.*

In sum, participants having *predictor conception* as dominant referred to this conception for different task scenarios and 1st summary question. Particularly, they defined the line of best fit as a line that gives closer predictions for values not included in dataset. In this respect, they assumed that there is an association between variables to be able to make predictions for a variable using the other. In parallel, they thought that the stronger the relationship would be resulted in the closer predictions. In addition, the participants also focused on placing the best fit line as close as possible to get more accurate predictions. In this respect, they showed similarity in criteria for placement to some extent with the participants having *representer conception*. However, in contrast to the participants having *representer conception* that evaluated the best fit line as providing estimations for where the data points are situated, teacher educators having the *predictor conception* utilized the line of best fit beyond the given dataset with aim of making predictions for the future and unknown values of the variables.

Signal: There are two teacher educators (TE9 and TE10) reflecting *signal conception* on their statements dominantly. In addition to them, TE2 also frequently referred to this conception. These participants generally interpreted data as a combination of linear structure and variability around it. In order to give better descriptions for the characteristics of signal conception of the line of best fit, I firstly share example excerpts from statements of each teacher educator who frequently referred to this conception throughout the interview. Then, I summarize the meaning of signal conception by using these teacher educators' perspectives.

For Task 1, TE9 made explanation based on the signal conception. In particular, he stated that:

"If the expected point were, x=10.8, let's say according to my line, it would be maybe a lower ... *not 6.1 but 5.9. The n here is an error. If we add up the errors, we know that we will find zero.* For instance, I thought to have minimum value of sum of the squares. That's why *I passed it through there so the errors would balance each other.* If I hold it up, the error squares of the points below the line will be very high. Since I thought that, I did something like this..."

As the data indicated, he mentioned the measurement error between the estimate value on the best fit line and actual data. In particular, he initially discussed deviation of an individual value on the best fit line from the actual data point and then thought taking average of the values to provide balance of residuals on each side of the best fit line. Therefore, he appreciated the variability in data set while placing the line of best fit. In addition, his statement as "*...If we add up the errors, we know that we will find zero…"* indicates that he aimed elimination of residuals while determining the line of best fit. This idea is similar to the signal-noise metaphor that defined the best fit line as a signal after eliminating the inherent noise. Therefore, I categorized this explanation as the *signal conception.*

Particularly, For Task 3, TE10 depicted signal conception as follows:

"We are looking for linearity and *there is no linearity that is normal and social in all reason.* No factor in bilateral relations… *let's assume we eliminate all. We wonder what kind of a relationship exists among these, it's actually an assumption.* These are totally about an assumption, *can I get a relationship among these if I rule out the other variables?,* If I could how would it be? *it would be linear, it would be a factor that affects one to one. However, since here other variables also affect, and since the point we focus on is two variables, we are trying to eliminate other variables and minimize the errors."*

As data showed, he stated that there is no perfect linear relationship in normal and social situations. In fact, there were some affecting factors on association between two variables. Therefore, irrelevant information might be included in the analysis process of bivariate data and cause some deviations from perfect linearity. Hence, he thought to eliminate inherent noise arose with other factors and get an approximation for the best fit line that shows relationship between given two variables. Then, the resulting regression line would be a *signal* for the association between given variables.

In addition, TE2 used the signal conception while elaborating on Task 5:

"….Now, it's harder to find the best fit line. Probably, what you are looking for is this, because in the data there isn't a trend like in the others (tasks). I shall not say there is no trend but there is less correlation between two variables. *When we look at the data, the others were more linear, but the residual values in the line of best fit here will be different.* It is almost impossible to find the line of best fit from it. *While the others have more correct line of best fit, they give more or less a more correct line of best fit, error probability of the one I'll give here is really high***…**But I will think in the same way and say something like this but my reliability here is quite low compared to the other ones, *there is higher deviation of this line."*

As excerpt indicated, since the dataset on Task 5 spread out more than previous tasks, correlation between variables would be lower. This means that residuals would have higher values when he determined the best fit line. According to *signal* conception, each value on the best fit line could be "viewed as deviating from the actual data by a measurement error and the average of these values could be interpreted as a close approximation to the actual data" (Konold and Pollatsek, 2002, p.269). With reference to this conception, although he could reach closer approximations when he placed best fit line

on prior task conditions, he interpreted Task 5 dataset including more deviations from line of best fit. In this process, he used the signal meaning of the line of best fit.

Lastly, TE2 and TE9 referred to the signal conception while answering the summary question. In particular, TE2 stated that:

"The notion to extract linearity from a data that is not linear is a difficult issue for them (students) to understand. *What we try to do when we explain the line of best fit is to explain that we are looking for an order in the mess of the data and that one of the best way to express this search for an order is line…"*

Given excerpt suggested that, he mentioned that students experience difficulty while finding a linear trend in dataset including variations from perfect linearity. In particular, TE2 addressed to this difficulty in explanations relevant to aim of this process as extracting the line of best fit that is showing and determining the order/stability in the data varied to some extent. This indicates his use of *signal* conception.

After that, TE2 further explained his reasoning on the line of best fit that reflects signal conception as follows:

*"…when we pass to line, the line gives a value…but we want to make interpretations for actual data values from a value on the line. We have an error margin***,** it might be hard to understand this…When we were making interpretations about this, we actually created the line of best fit, we have the data points in front of us. To see where the each data point would be on the best fit line…I mean, if I did not see the data point, what I will predict with this value, the best fit line...But the actual value is this. Or I would predict a higher value as a lower value. Or there is a point that fits, I would predict that. I show these three situations on a data and then I predict a point that I don't know. Therefore, one of these three situations is possible. I mean, *after explaining the concrete situation, predicting the abstract thing that doesn't exist. Because we don't see these (data points) anymore. Then, after extracting actual data points, my predicted values are there, but they actually might be around there as we see before. Those things which we call residual, the margin of error, show us the places where this data exist and it shows it with some error and we can say that it gives us that this margin of error exists."*

As the data showed, he defined the line of best fit as a structure extracted from the data points varying in the data set. This was evident in his statement "*when we pass to line, the line gives a value…but we want to make interpretations for actual data values from a value on the line. We have an error margin...*" So, he conceptualized the line of best fit independently of the data points---the signal of the data---since he aimed to reduce the inherent variability to reach the stability within the process varying to some extent. He also stated that students need to be thinking the line of best as the *signal* of data as well as thinking the actual data values and the residuals in the data set. That is, from his point of view, students need to be focusing on the line of best fit as a linear structure that gives predictions for variables while taking into consideration the margin of error in these predicted values and point to the three cases: that the values on the line could be the same as, smaller or higher than the actual data value. This excerpt showed that he used *signal* conception.

In addition, TE9 also explained summary question based on signal conception:

"*…We make them calculate the residuals. We let them see this and that it's different for each person.* For instance Ali got 50 from the exam, the expected value is his getting 60, that is the difference in between. *We show them the value and the errors. We say that the value which makes the residuals minimum but we also say that it can be calculated for each person if necessary."*

As data suggested, he mentioned each individual value on the best fit line as an estimate value deviating from the actual data with some error. Particularly, he interpreted the dataset as a composition of the *signal* that gives close approximations for data values and *deviations* from actual data points. This leads to *signal conception* of the line of best fit.

In sum, participants having *signal conception* as dominant referred to this conception for different task scenarios and 1st summary question. In particular, they interpreted the dataset as a combination of line of best fit and residuals around it (data=signal+noise). This perspective is useful to construct a bridge between deterministic and stochastic approaches.

In general, they believed that perfect linear association does not occur in real life situations. For them, there might be other affecting variables that cause some deviations from linearity. Therefore, each individual value on the best fit line has some deviation from actual data by a measurement error. Hence, they thought to eliminate inherent noise arose with other factors to deal with relevant and irrelevant information. Therefore, they aimed to minimize errors and provide balance of deviations from data values. For this conception, teacher educators evaluated task conditions based on the data points given on scatterplot, the line of best fit that closely approximates the actual data values and also residuals that are emerged with the effect of other variables in situation simultaneously. In particular, since they analyzed each individual value and also the dataset as a whole concurrently to find the best fit line, evidences for the *representer conception* might also be captured from these statements. In addition, since they aimed to get close approximations to the actual values and made reference to functional structure as the regression equation, I could say that these participants also reflected the *predictor conception.* Therefore, *signal conception* might be thought of embracing the other dominant conceptions as *representer* and *predictor*.

After all, teacher educators having the *representer conception*, are focused on visual representation of given data set at hand while placing the informal line of best fit. On the other hand, while the teacher educators with the *representer conception* as their dominant conception had the intention to determine how the data accounts for the relationship depicted in the particular scatterplot, the teacher educators having the *predictor conception* aimed at explaining the use of the line of best fit. That is, *predictor conception* allowed to focusing more on the points beyond the particular case such that the utility of the line of best fit is to provide predictions for unknown and future values with minimum errors from the actual points. Finally, *signal conception* has the conceptual basis that involves statistical perspective for meaning of linear regression. The participants having this conception searched for a linear structure in dataset including variability to some extent to be able to make an interpretation for the association between variables. In addition, they also dealt with hypothetically existed confounding variables having an impact on the association and creating deviations from linear association. Therefore, these participants defined the line of best fit as a signal after elimination of these deviations. In parallel, they

also appreciated the margin of error while interpreting the dataset. This indicated that they were more aware of the nature of statistical perspective for the line of best fit.

Table 5.2. Teacher educators' dominant conceptions including operational definitions, illustrative excerpts and frequency.

Conception	**Definitions**	**Example Excerpts**	**f**
Representer: TE1, TE4, TE6, TE7, TE8	-the best representation of all sample data, where the line accounts for data at hand, rather than more generalized relationship (Casey, 2015b)	TE4-"It is *the line that best, most optimally, represents whole set of data*." (1st summary question) TE7-"*The word represent*…I shall first think about the keywords and then I will use it in a sentence. *Represent means representing the data***.** Of course you can look at the context to see if it fits *but it's important that it represents the data.* While representing, deviation…I would say that is the balance of the distance of the points that are above the line and below the line. *So I can represent it.* Trying to balance a deviation that will stay close to the points."(1st summary question)	5
Predictor: TE2, TE3, TE5,TE11	-the line that enables predictions for values not in the dataset (Casey, 2015b).	TE2-"*We are trying to deduce a prediction from the data*. When we talk about the data at hand, we can see the results *but we are trying to make predictions about the data that don't exist.* In order to do this, we generally assume the data as linear, we try to fit in a linear model. We need to draw a line in the linear model that we are trying to settle. We need equations, functions for the operation that we call prediction and regression. We are trying to represent these functions with a line because it is easier to do so. *We are looking for the line which fits the data and which produces the least errors. That is the line of best fit***…**"(1st summary question) TE5-"*This line can be used to make predictions***.** For instance, if the independent variable is this, what would be the expected dependent variable value? *This line will make it is easier for us to predict this."(*1st summary question)	4

Table 5.2. Teacher educators' dominant conceptions including operational definitions, illustrative excerpts and frequency (cont.).

Conception	Definitions	Example Excerpts	f
Signal: TE9,TE10 (TE2)	-Thinking of each individual value on best fit line as an estimate of a specific value such that each value "is viewed as deviating from the actual data by a measurement error and the average of these values is interpreted as a close approximation to the actual data (Konold and Pollatsek, 2002) -Stability in variable system (Konold and Pollatsek, 2002) -the meaningful signal for data that eliminates inherent noise (Casey, 2015b)	TE9-"…We also make them calculate the residuals. We let them see this and that *it's different for each person. For instance Ali got 50 from the exam, the estimate value is his getting 60, that is the difference in between. We show them the value and the errors. We say that the value which makes the residuals minimum* but we also say that it can be calculated for each person if necessary." (1st summary question) TE2-"In other words, they (students) have difficulty in this: *The notion to extract a linearity from a data that is not linear* is a difficult issue for them to understand. *What we try to do when we explain the line of best fit is to explain that we are looking for an order in the mess of the data and that one of the best way to express this search for an order is line*…" (summary question) TE10-"We are looking for linearity and there is no linearity that is normal and social in all reason. No factor in bilateral relations… *let's assume we eliminate all. We wonder what kind of a relationship exists among these, it's actually an assumption.* These are totally about an assumption, *can I get a relationship among these if I rule out the other variable*s?, If I could how would it be? it would be linear, it would be a factor that affects one to one. However, *since here other variables also affect, and since the point we focus on is two variables, we are trying to eliminate other variables and minimize the errors.*" (Task 3)	2

Considering all these dominant conceptions demonstrated on Table 5.2 injuxtaposition to each other, while *representer conception* of the informal line of best fit allows visually focusing on the data points particular to the situation at hand, the *predictor conception* allows to consider the predicted values that are possibly on the line of best fit. Contrarily, the *signal conception* allows simultaneously and holistically to focus on the data points particular to the situation, the predicted values and the residuals.

5.2. Criteria and Methods

In this part, I aimed to describe teacher educators' criteria for placement of the line of best fit. Therefore, I analyzed teacher educators' statements from three data sources: 1) line

placement criteria for seven task situations with different contexts, scatters of data points and strength of correlation, 2) interpretations for Task 8 including students' line placements, 3) 2nd summary question as "Could you tell me what you would say to a student to help them draw the line of best fit on a scatter plot?" In the analysis of these data, I firstly determined criteria categories that teacher educators referred throughout the interview process. Then, I identified dominant criterion for each teacher educator according to the frequency of their reference to these criteria categories.

In the analysis of teacher educators' descriptions for their placement criteria of line of best fit, I identified six different criteria, namely, *closest, sum deviation, reflecting context, ellipse& area approach, equal number and pairs*. These criteria categories and the descriptions relevant to them are thought as referring to the research question as "What criteria do teacher educators use for placement of informal line of best fit?" Each teacher educator has more than one criterion to place an informal line of best fit on different task situations. In particular, *closest* criterion used by all teacher educators at least once, *sum deviation* criterion applied by eight teacher educators (TE1, TE2, TE3, TE4, TE7, TE8, TE9 and TE10), *reflecting context* criterion used by four teacher educators (TE1, TE6, TE9 and TE11), ellipse and area approach referred by two teacher educators (TE2 and TE4), equal number criterion applied by three teacher educators (TE2, TE3 and TE10), and pairs criterion is utilized by two teacher educators (TE9 and TE10) while placing the informal line of best fit. Below, I constructed Table 5.2 representing six criteria categories held by teacher educators, their operational definitions, illustrative excerpts for each category and frequency of participants who used them to place the informal line of best fit.

Table 5.3. Criteria categories held by teacher educators with operational definitions, illustrative excerpts and frequency.

Criteria	**Operational Definitions**	**Example Excerpt**	**f**
Closest All participants	The line that is closest to all points (implies total distances from data points	TE5-"So I tried to draw *the line which is close to all points. Because it is in the logic of the best fit, in order for it to be the best fitting line, we need to minimize the margin of error*. In order to minimize the error ... *I tried to minimize the error*	11

Table 5.3. Criteria categories held by teacher educators with operational definitions, illustrative excerpts and frequency (cont.).

Criteria	**Operational Definitions**	**Example Excerpt**	**f**
	to line is minimized)	*as it is related to the distance between these points and the line (drawing vertical line between the point and the line) and I tried to make it closer* to the area where the maximum number of points exist." (Task1) TE11-"As far as possible, *I should choose a line which is the closest to these points, so that the distance between these points and the line would be the smallest. I can say that…So, we have made this decision considering the smallest distances*." (Task1)	
Sum deviation TE1, TE2, TE3, TE4, TE7, TE8, TE9, TE10,	Sum of deviations for points above the line is the same as the sum of deviations for points below the line (they did not necessarily make measurements, but referenced this as they placed their line)	TE1-"Actually, after drawing this line, *the distance between points and the line seemed balanced for both sides.* So, we are calculating r, correlation coefficient. This coefficient is some way related to these distances. If all of the distances are zero, then r becomes 1 or -1. *As the distance increases, the balance of the distances from points on the right and left to the line might be also important;* I don't remember the exact formula." (Task 4) TE3-"If I did my best while I was placing these lines, *the sum of the distances would already be zero*. That's the underlying process of the line of best fit. For example, if I claim that this line is the line of best fit in Task1 or 2, I have to come up to the conclusion such that *the sum of all distances should be zero*. I hope they (the sum of the distances) are zero or close to zero. Speaking of distance, we call it $y-y_{predicted}$ and we do not change it at all. *We get negative values from the differences below the line and positive values for the points above the line. So it is not the distance, it is the sum of residuals. Hence, the sum of residuals should be zero.* The assumption of the line of best fit is actually this. *If it is zero, our line is really the line of best fit. If it's not zero, there is a better line than I've drawn."* (Task2)	8
Reflecting Context TE1, TE6, TE9, TE11	The line that is reflecting the relationship the variables have based on the context knowledge (Casey, 2015a).	TE1- ".. *regarding the context, a taller person's shoe size is expected to be a little longer, isn't it? I mean, it is a slightly more positive relationship…* So what is meant in the context is not always consistent with what the data says. Sure, there are 8 people here. *That is, it doesn't have to be so, but it influences one's decision... If you do not show me any graphs, I would say one increases as the other increases; I would say there might be a positive relationship. Right here the line is someway similar to this.* But I think we can not identify a clear linear trend for this data set." (Task 6) TE6- "It says primary school teacher. I mean, one will be 4th grade. How old is the child in forth grade? 10 years. *The child doesn't wear size 7 shoes when he is 10. The size he can wear is no more than 5 or 6, I think it can not be more than that.* When I examine it, I already see that the data is gathered in this area. *So I believe the line should be like that. That is, the person who has size 2 shoes is 50 cm tall, the other who has size 3.5 shoes is 51 cm tall... After all, I can say that there is a linear relationship between the shoe sizes and the students' height.*" (Task6) TE11-"Now, of course, while drawing the line, I'm going to reason it out. *Is there a linear relationship, for example, such as his shoe size gets larger while the child*	4

Table 5.3. Criteria categories held by teacher educators with operational definitions, illustrative excerpts and frequency (cont.).

Criteria	Operational Definitions	Example Excerpt	f
		gets taller? Because it shows not much of a relationship as it is dispersed. If we start to think whether it is in a positive or negative direction… I feel like, just by looking at it, it is much better (drawing a positively sloped line for example)... *So I asked what could be the relationship between the height of a person and the size of the shoe; I thought as the child's height increases, the size will also get larger. Because younger kids have smaller feet."* (Task 6)	
Ellipse & Area approach TE2 (for weaker association) TE4	Drawing an ellipse or an area surrounding to the all data points and determining the line by dividing it symmetrically.	TE2-"We are trying to see the trend, what we call 'the trend' in dataset is the relationship between variables as the increasing-increasing, decreasing-decreasing, increasing-decreasing, I mean at first we are trying to see what is the increase or decrease amount corresponding to an increase. While determining this trend, I am trying to see how much of this data on the linearity. *We also draw an ellipse including the data points to determine the correlation. The weaker and thinner the ellipse is, the more likely it is to show us that the line exists here. So, the logic behind drawing an ellipse is that should show me two points through which line passes (showing the points at the two endpoints of the ellipse). So as the line passes through two points, I'm actually trying to find where those two points could be. The ellipse slightly shows me their places. It is the reason I draw this ellipse.* The problem starts when the ellipse starts to get rounded. When the ellipse becomes too round, you can not be sure where those endpoints are. So when the dataset is too dispersed, finding the endpoint is getting difficult. Our problem starts here. *But if the ellipse is weaker, it helps me determine the endpoints."* (Summary Question -ellipse) TE4-"I was drawing like this for this case. *I consider it as a piece of wood, and I'm trying to find a line that will divide it into two equal parts.* While doing so, let me say that I could have drawn it in the concave form, but I consciously used the convex form. In other words, from the psychological point of view, I wanted to reduce multiple data without losing their values. For example, *I imagined a piece of wood like that. While I was searching for a line that would divide this piece of wood into two,* I aimed to go beyond these data points and use them without losing the values they represent.. *So I created an area here, actually, we can call it as the area approach, so I divided this area into two pieces. I mean, the method as dividing the wood into two pieces is easier than considering all points at the same time"* (Task 1 -Area approach)	2
Equal number TE2, TE3, TE10	Equal number of points above and below the line	TE2- "But what I'm trying to do here while choosing the best fitting line: once I want it to pass through all points, but because I'm sure it will not be so, *what I'm trying to achieve is to provide having an equal number of points* and equal distances *above and below the line,* in other words, it is what we call the best fit line: I try to minimize the residuals of this regression, I mean minimization of errors. *While trying to minimize it, I try to keep the number of points as equal as possible above and below the line. Approximately 3-4 data points are above, 3-4*	3

Table 5.3. Criteria categories held by teacher educators with operational definitions, illustrative excerpts and frequency (cont.).

Criteria	**Operational Definitions**	**Example Excerpt**	**f**
		data points are below the line. But as some have more residuals, the number of points above the line will be relatively more. While determining this best fitting line, we did not give decision regarding the context but rather we decided it with consideration of the residuals." (Task 1) TE3-"I have constructed a weak, almost zero, and positive relationship. Then, *I constructed my line of best fit as approximately 4 points above and 4 of them below...*" (Task 6) TE10- "Here, too, I guess *we will probably obtain a line which centers these points.* You will ask why.. because there are *3 points, deviation, and there are also 3 here, if you think them as up and down, we would balance them. In other words, if I can create a line with the distribution of 3 points above and 3 below,* I guess it will provide linearity." (Task 4)	
Pairs TE9, TE10	Each point above the line has a symmetric pair below the line so that the distance to the line is the same for each member of the pair; pairing off individual data points to determine line of best fit	TE9- "Here again *I can try to center this line in order to eliminate the errors. Let me think of the errors ... in twos ...I mean I saw in twos like that, I considered the two of them (the last two data points), I thought of the two of them together (delineating) and I thought these like symmetrical to each other, I mean, this one and that one, in order to balance*.."(Task 3) TE10- "Because *one is plus and the other is minus,* look… *the values are exactly close to each other and the center is just like...* In other words, *I can easily draw the exact line which passes through just between these two points."* (Task 2)	2

After that, I also determined dominant criterion for each teacher educator based on the frequency of their use of the criterion for placement of informal line of best fit. In particular, three criteria categories were depicted as dominant criterion held by teacher educators: *closest, sum deviation* and *ellipse and area approach.* In particular, seven teacher educators (TE1, TE2*TE3, TE5, TE6, TE8* and TE11) used *closest* criteria dominantly, five teacher educators (TE2*, TE7, TE8*, TE9 and TE10) applied *sum deviation* criterion dominantly and two teacher educators (TE2* and TE4) used *ellipse and area approach* criterion dominantly. Besides these, TE2 and TE8 used different criterion categories with the same frequency. Therefore, they are included in each criterion categories they used dominantly. Since the focus of this study is also to describe characteristics of criteria held by teacher educators dominantly, in the following paragraphs; I elaborate on the data regarding dominant criteria teacher educators depicted.

Closest criterion: Closest criterion was referred by each teacher educator at least once during line placement process. In addition, seven teacher educators (TE1, TE2*, TE3, TE5, TE6, TE8*, TE11) utilized it dominantly to determine the best fitting line on different task situations. In respect to understanding characteristics of closest criterion for placement of the informal line of best fit, I firstly share example excerpts from statements of these teacher educators who dominantly used closest criterion throughout the interview. Then, I summarized their statements to describe the closest criterion better.

For Task1, I share example excerpts from statements of TE3, TE5 and TE8 focusing on the closest criterion while placing the best fit line. In particular, TE3 stated that:

"…*we are trying to minimize the sum of the squares.* The line of best fit is based on this structure. That's what they call 'the least square regression method'. *Keep the distances as less as possible, minimize the squares***.** I've paid attention to this in general. *What I mean by distance is, let's say the coordinates of the data point is actual x or actual y, the coordinates on the line are their predicted values. I minimize the distance between them (coordinates of actual and predicted x and y values),* so that my model, the model that I will obtain by drawing this line segment will give the results that are close to the actual data values as possible. x or y does not matter, y-$y_{predicted}$ *is the residual for each one (data point) there. We're trying to minimize this."*

As excerpt showed, he explained the line placement process based on *closest criterion* referring to the least square regression method. Therefore, he aimed to provide minimization of square of residuals. In particular, he defined residuals as difference between actual data values and their predicted values on the best fit line. Hence, he aimed to reach minimum value of this difference to get the line of best fit. Thus, he could make closer approximations to actual data values.

Similarly, TE5 mentioned that:

"Now we can draw the best fitting line in this scatterplot as that, *as for the best fit line, the distance between the points and the line needs to be least.* If we do it like this, we're approaching the second point, but we're moving further away from the above points. *Since the number of the points is higher there (showing the last 4 points), I wanted to be closer to them. So I tried to draw the line which is close to all points.* Because it is in the logic of the best fit, *in order for it to be the best fitting line, we need to minimize the*

margin of error. In order to minimize the error ... I tried to minimize the error as it is related to the distance between these points and the line (drawing vertical line between the point and the line) and I tried to make it closer to the area where the maximum number of points exist."

As data suggested, he mentioned that the vertical distances between data points and the line should be minimized to provide the best fitting line for given dataset. In particular, he tried to be closest to all data points while placing the line. In this way, he applied the closest criterion on this task.

Also, TE8 described his line placement process on Task 1 focusing on closest criterion:

"In other words, it is essential that the *distances from the points to the line are minimized.* At the moment, *I think that it will be the minimum* just by looking at it. So I think this might be the line that represents them. Or a more special... for example, *I am not affected by the context. The situation would be the same even if it was something other than a ball. So here it is important that this line cannot be horizontal, cannot be vertical, it is essential that the vertical distances of these points are minimized*. I can think of something else as well, we can also directly use the distances between the points and the line. I know such a study. It is called diagonal regression. It has longer calculations but when you compare the two it (diagonal regression) gives a line with less error. There are very few differences. It is quite close to each other but *in traditional literature it is essential that the distances on the y axis are minimized; which are known as the sum of* e^2. I think this line is the best line to represent this data set."

As data suggested, he mentioned the minimization of distances from data points to the line as an essential approach for the line placement. In addition, he emphasized that he focused on the scatterplot of given dataset instead of reflecting the relationship based on given contextual information. Additionally, he mentioned about a research study relevant to the linear regression and least square regression method. According to this study, he referred to the calculation of residuals and regression equation in two ways: using vertical distances (least squares regression method) or finding perpendicular distances (they called as diagonal regression) between actual data and the best fit line. In particular, results showed that making calculations using the perpendicular distances give lower margin of error than finding residuals vertically. However, since the difference was not significant

and also procedures were more complicated for diagonal regression, the procedures of least squares regression method are applied mostly.

In addition, TE1 and TE11 demonstrated closest criterion on Task 5 including more dispersed dataset and also an outlier point (2, 100). Particularly, TE1 explained as:

"… I drew the line according to the *other points regardless of that one. No matter what I do, this point (2,100) would still be far from the others. If I try to make closer to it, the line would get further away from other points.* In other words, since there was an accumulation on this part (pointing with his hand), *I tried to draw the closest line to the other points without considering the extreme point."*

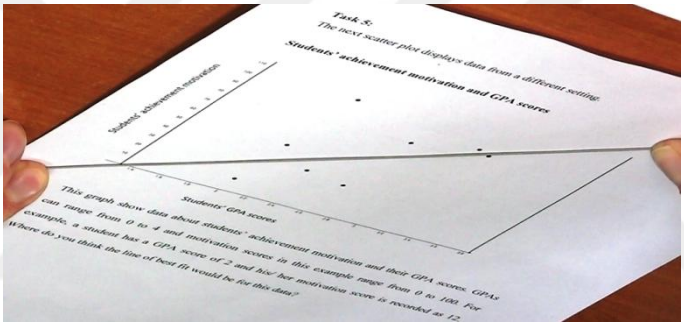

Figure 5. 1. TE1's line placement on Task 5.

As data and Figure 5.1 suggested, TE1 determined the best fit line on Task 5 by applying the closest criterion for seven data points. In particular, since there were an extreme score (2, 100) that is far away from general trend of dataset, she preferred to discard this point. In this respect, she thought that if she includes the extreme score to the line placement process, it would cause an increase on total distances between data points and the line. Hence, with the aim of minimizing sum of residuals, she used the *closest* criterion for seven data points out of (2, 100).

Similarly, TE11 interpreted this task as follows:

"Yes, it's a bit more dispersed, so *if you want to make this distance as less as possible… for example if I draw it closer to this side (to the lower part of the graph), this time they are more distant. So… if I draw it closer the other side (to the upper part of the graph), these will become distant. I tried to determine that the distances to the line are as small as possible ."*

As excerpt showed, she focused on providing minimum distances from data points to the line. Since scatterplot of data points seemed more dispersed on Task 5, she was more irresolute compared to previous tasks. However, he again put the informal line of best fit with aim of having closest distance to all data points. Contrary to TE1, she handled dataset as a whole including the extreme point (2, 100).

By the same token, while elaborating on students' line placements given on Task 8, TE2 and TE6 made evaluation based on closest criterion for placement of informal line of best fit. Therefore, TE2 interpreted two students' line placements on Task 3 as:

"*The student is probably trying to pay attention to the residuals* with similar reasoning, but these (the residuals above and below the line) almost all eliminate each other, but *when we look at the total residuals, Jane has more than Angelo*. It is not only trying to eliminate these residuals but also *we are trying to minimize them. Angelo has the situation where we try to achieve the minimum.."*

As data suggested, he evaluated the residuals between data points and the line for each student. For him, both students provided the elimination of residuals above and below approximately. However, what is more important is that achieving the minimum distances in total. Therefore, he determined Angelo's line placement as the best since it was closest to all data points.

In parallel, TE6 interpreted students' line placements for Task1 referring to the closest criterion:

"Angelo makes it pass through at least two points, in fact, *Barbara drew the closest line. That's why I think Barbara's drawing is more accurate."*

As excerpt suggested, she made comparison of two students' line placements for Task1. For her, while Angelo was placing the informal line of best fit as passing through two data points, Barbara gave importance to minimize distances between data points and the line. In particular, since Barbara's line seemed to be determined as close as possible to the data points, TE6 preferred her placement as better than Angelo's.

Lastly, teacher educators also referred the *closest criterion* for 2nd summary question "Could you tell me what you would say to a student to help them draw the line of best fit on a scatter plot?" as follows:

In particular, TE1 stated:

"The method that I have applied just now is that**:** *the line that is the closest to the points is the right one. I just need to think about all the points, not just some. Of course there may be some extreme values, but I try to draw a line that passes near the majority of all these points.* I mean I will not draw the line through some of these points... the thing is: *two points are enough to draw the line graph, all we do is to find two points and draw them together. But there's no such thing. These are non-linear anyway*, we already realize that. *Drawing a line passing through either of them is useless*. So I'm not trying to draw through any point, rather *I'm trying to draw a line close to all points***.** Because of that**, I** *need to check the distance between the line and the points."*

As data showed, TE1 summarized her dominant criterion as determining the line that is closest to all data points. By doing so, she stated that she focused on the whole data as opposed to any particular data point. In addition, she referred to the criterion as 'passing through two data points' which is thought as one of the most common misconception for students. Particularly, she stated that the regression line obtained with this criterion would not be useful if the dataset is varying to some extent. Therefore, she underlined that she did not determine the best fit line as passing through any data point intentionally, instead, she aimed to place the line of best fit as closest to all data points.

Also, TE3 explained this question referring to the closest criterion as:

"Place such a line of which you *try to minimize the vertical distances between each point and the line you draw with eyeballing*. But, consider the sum of them each time. *Try to minimize the distances both for each point and for sum of them as well*…"

As data indicated, he described the line placement process as minimization of vertical distances between each data point and the line. In addition, he also pointed to the consideration of minimizing sum of residuals at the same time. This approach reflected his criterion for placement as the closest.

Similarly, TE5 concluded that:

"What we call best fit here is that *we have to choose a line that has optimum closeness to all points***.** So you can prefer a line which passes through a few points, but this may not be the best fit. Hence**,** *we have to make a choice for a line which has the optimum and shortest distance to all points...."*

As data indicated, he described the best fit line as having optimum distance from data points. Similar to TE1, he also mentioned the criterion as placing a line passing through some of the data points. Particularly, he stated that this method would not result in the best fit line for given dataset. Therefore, he suggested to using the closest criterion in determination of the line of best fit for dataset including variation.

By the same token, TE8 answered the summary question based on the closest criterion with explanations relevant to the least squares regression method. So, he stated:

"…square of errors… *I thought the minimization of the sum of distances***.** Then I start to think the sum of distances will not give a result since some distances are positive and some of them are negative... but *it can be obtained from the sum of the absolute value of distances*…in fact, as in standard deviation or absolute deviation calculation, *the value obtained from the sum of the squares of the applied distances is even more reassuring* or is at least a scientific value…"

As the excerpt indicated, he thought the minimization of the sum of distances between data points and the line. However, since the distances could have positive or

negative values, he mentioned minimizing sum of squares of distances. In this way, he referred to the deterministic formula to find the best fitting line as least square regression method. Hence, he used the closest criterion while explaining the reasoning on line placement referring to the procedural methods.

Lastly, TE11 explained as:

"I would explain to him like this: I mean, for example you will try to guess something, like, the child studied 3 hours and got 10 but studied 4 hours and got 7.. I would give an example of a data group like this. Then I would ask if there is a relationship between them.Let's try to decide considering the available data. *We examine their distance from each other. We would give one sample data, the child could approach the right value,* or for each one, I would say that the child studied for 8 hours and got 13. *Try whether it also gives the closer values,* then the child could find 13.1 and write down the answer and find 13.5 for the other... *We want him to find the closest result by substituting on the linear equations.* So which one help me make the predictions easier is where *these points are at the closest distance both to each other and to the line...."*

As the data showed, she started to explain with an example of hypothetical dataset regarding the association between students' study hours and their exam grades. Then, she assumed leading students to examine the scatterplot of data and determine the line of best fit based on it. In this case, her focus was to minimize the distances from actual data values. Therefore, she thought to suggest students to substitude some values to make comparison between actual data and predicted values after they placed their lines. Hence, she concluded that the informal line of best fit would be the line providing minimum distances between predicted values and actual data values. Therefore, she suggested the use of closest criterion for placement of line of best fit.

In sum, the participants having the *closest criterion* as dominant aimed to place the informal line of best fit as close as possible to the data points on scatterplot. In this respect, they handled data points as a whole. Particularly, they did not have tendency to place a line passing through some data points, instead, they focused on minimizing total residuals from all data points. In addition, the procedural method of finding the line of best fit as least square regression is also based on minimizing the sum of square of residuals. Hence, the

closest criterion for informal line of best fit goes along with the calculational methods for linear regression. Therefore, the participants using this criterion dominantly for placement of informal line of best fit have the general knowledge of underlying procedures for formal regression.

Sum Deviation Criterion: The sum deviation criterion was obtained from five teacher educators (TE2*, TE7, TE8*, TE9, TE10) dominantly throughout their interview process. In order to describe characteristics of the sum deviation criterion for placement of the line of best fit, I firstly present illustrative excerpts from statements of teacher educators who dominantly referred to this criterion. Then, I elicit the procedures of the sum deviation criterion based on teacher educators' perspectives.

To begin with, TE7 described her placement criterion for Task 1 and 2 based on the sum deviation as follows:

"Rather than minimizing the distance in both, *my strategy was to keep the deviation under control (in task 1)… to balance the distances above and below…* So in fact I didn't try to minimize the distances. I have to say this. I'm always examining the distances. But rather than minimization of the distance .. I even told you that it didn't even have to pass through a point. What is the minimum distance? It is zero. If I try to minimize the distances,I need to find the line passing through the majority of points, but I'm not making such an effort**.** *I'm always considering the deviation. In both of them."*

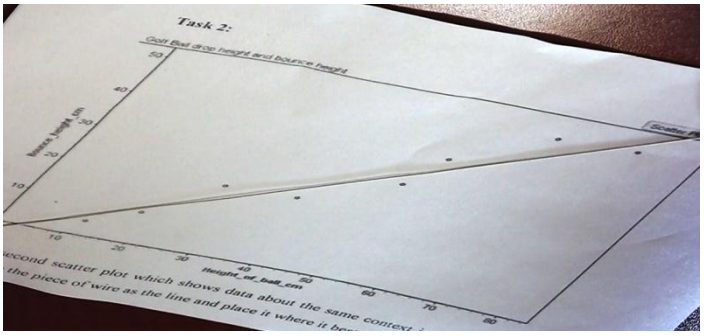

Figure 5. 2. TE7's line placement for Task2.

As data and Figure 5.2 demonsrated, she focused on controlling deviations while placing the best fit line on Task 1 and 2. In this respect, she aimed to balance the sum of deviations above and below the line. In addition, she stated that she did not try to minimize distances between data points and the best fit line anymore. For her, distance minimization could be provided by placing a line passing through most of the data points since she interpreted the minimization as reaching to zero in distance. Hence, her reasoning while placing line is based on *the sum deviation criterion* dominantly on these tasks.

In addition, for Task2, TE9 also explained sum deviation criterion:

"Again, *I'm trying to center a little. I thought that if it passes through just one, and the error of another one is 2 and the other is 0; it would give an error of 4 units. If it is 1, 1, 2 units will be the result. So it's reasonable to balance them. So I tried to center it approximately,* but there are three data points and here are five when it is passed from center. So I've tried to get a little bit closer to five of them, not so much, but a little bit…for could it be 4 to 4? It appears to be millimetric, to me..."

As the data suggested, he placed the informal line of best fit to the center of data points as balancing distances on each side of the line approximately. In particular, he thought that if he preferred to place a line passing through some data points, these data points would have zero deviation from the line. However, the residuals might have higher values for other data points. As an example, he compared two conditions: r_1=0 and r_2= 2 or r_1=1 and r_2=1 as residuals for two data points on different sides of the line. In particular, if one calculated the square of these residuals based on least square regression method to find the best fit line, the second situation would result in minimum values. Therefore, he concluded his placement criterion as balancing the deviation on each side, instead of placing a line passing through some data points.

By the same token, TE2 referred to this criterion on Task 3:

"There is a linear association here. Examining in a similar way, *what I'm trying to calculate is that: I'm trying to keep the residuals above and below the line approximately the same.* I suppose it would be something just like this."

As data pointed, he stated his placement criteria such that he aimed to keep residuals above and below the line approximately the same. This indicates his sum deviation criterion for this task.

Similarly, TE8 elaborated on Task 3:

"...I left the four points on one side, this time it is equal number, *I left the four points but the numerical equation is not an essential parameter to place the line. The important point is the balance of the sum of distances on each side.* In fact, the way we place it now, the distances above the line were more effective, maybe. So we can move it a little more."

As excerpts suggested, he aimed to provide the balance of sum of deviations above and below the line. For this aim, he firstly controlled the number of data points. However, he further explained that it was not necessary to satisfy the equality of number of data points, rather he focused on the equivalence of sum of deviations on each side of the informal line of best fit.

Also, TE10 placed his line on Task 4 using the sum deviation criterion:

"Here, too, I guess *we will probably obtain a line which centers these points.* You will ask why.. *because there are 3 points, deviation, and there are also 3 here, if you think them as up and down, we would balance them.* In other words, if I can create a line with the distribution of 3 points above and 3 below, I guess it will provide linearity."

Data pointed that, he aimed to obtain a line that centralize the total distances above and below the line. In this process, he placed the informal line of best fit as passing through two points and having three points on each side of the line. By equating the number of data points, he enabled the balance of residuals above and below the informal line of best fit.

Lastly, for 2nd summary question asking "Could you tell me what you would say to a student to help them draw the line of best fit on a scatterplot?", TE7 and TE9 referred to the sum deviation criterion. In particular, TE7 stated that:

"For the child to understand, I would say something like the *above distances and the below distances ... What do we mean by below and above? I would say to him that the answer is the balance of the distances between the above points and the line, and between the below points and the line.* So I can represent it ... What does it mean to represent? *It is to try to balance a deviation which not has incoherent distances, but is close to the data points."*

As excerpt indicated, she defined the line of best fit as a representation of given dataset. Therefore, to be able to reach the aim of representing dataset, she thought to lead students to determine the line of best fit by providing balance of distances between above points and the line, and below points and the line. Then, she further stated that the line of best fit should also be determined as close to data points as well as providing balance of distances on each side of the line. Therefore, although her dominant conception was the sum deviation criterion, she supported her perspective with closest criterion.

Similarly, TE9 mentioned:

"I would like to discuss what would cause more errors. *If there is an error of 3 units the effect will be 9, 2 unit, if we divide it 2 as 1 it would be equivalent to 4 to 1, so that the errors must balance each other…* but in order to help the students*… error concept, like in which conditions the error will be less, they should balance each other…* I don't want to say it should be centered. It's because it may not center in this situation. But *we've seen in most of them that it passes through the center.* It can also be useful to discuss the matter."

As data indicated, he made reasoning for criteria for line placement referring to the procedures of the least square regression method. Correspondingly, he thought providing minimum value for square of residuals to reach the best fitting line. In this respect, he aimed to provide balance of distances above and below the line. Then, he thought that when he balance the residuals, he could support the idea behind procedural methods and so find the informal line of best fit. Therefore, he suggested *the sum deviation* criterion for placement of informal line of best fit.

In sum, the participants using sum deviation criterion dominantly aimed to provide equivalence in sum of deviations above and below the informal line of best fit. In this respect, they described the deviations as distances between actual data values and the line of best fit. Then, they paid attention to balance of these distances on each side of the line. By the same token, some teacher educators also mentioned this criterion as averaging the sum of deviations while placing the line. In fact, as TE7 stated, if the sum deviation criterion as balancing deviations above and below the line and closest criterion as placing the line closest to all data points are used simultaneously, one could reach the informal line of best fit.

Ellipse & Area Approach Criterion: In this research, there is an emerging dominant criterion as ellipse and area approach in addition to the criteria categories mentioned in the previous studies with students and teachers (Casey, 2015a,b; Sorto *et al.*, 2011). Although two teacher educators (TE2 and TE4) classified in this category used different terms as "ellipse" and "area approach", I combined them since the method and perspectives showed similarities. In order to give better descriptions for the characteristics of ellipse and area approach for placement of the informal line of best fit, I firstly share example excerpts from statements of TE2 and TE4 who referred to this criterion. Then, I summarize these teacher educators' perspectives to extract the utility of this approach.

For Task 1, TE4 described his line placement process as follows:

"I was *drawing like this for this case. I consider it as a piece of wood, and I'm trying to find a line that will divide it into two equal parts.* While doing so, let me say that I could have drawn it in the concave form, but *I consciously used the convex form*. In other words, from the psychological point of view, I wanted to reduce multiple data without losing their values. For example, *I imagined a piece of wood like that. While I was searching for a line that would divide this piece of wood into two,* I aimed to go beyond these data points and use them without losing the values they represent. *So I created an area here, actually, we can call it as the area approach, so I divided this area into two pieces. I mean, the method as dividing the wood into two pieces is easier than considering all points at the same time".*

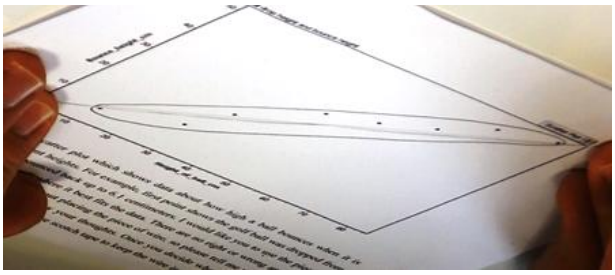

Figure 5. 3. TE4's line placement for Task 1.

As data and Figure 5.3 showed, he initially created an area including all data points. By doing this, he believed that he could handle the dataset as a whole. Then, he aimed to determine a line dividing the area into two equal parts. Hence, he called this criterion as the area approach. For him, this method as dividing the area into two equal pieces would be easier than considering all data points at the same time while placing the informal line of best fit.

On the other hand, TE2 applied similar criterion while placing a line for the tasks displaying data with lower associated or nonassociated variables. However, he used different term –ellipse-- while describing the process. In particular, for Task 5, he stated that:

"…if *you try to draw an ellipse around it, or to draw a round, in fact, this ellipse gives us the shape.* Now, taking the previous task as an example, *if we draw an ellipse here, the direction of this ellipse already gives it. So the ellipse is associated with the line.* But the thing I'm trying to draw here is that**…** *it seems like outlier ((2,100) point).* I use the expression 'like' as *I am not sure whether it is exactly an outlier or not.* But *considering this group, let's assume that I remove this point (the point assumed as an outlier),* let's put an eraser on it. *These show us that there is a trend, more or less, although it is weaker than the others. But when you add this data point, this point will naturally affect the relationship. I don't know how much it will change. But the ellipse which you draw when you want to include this point is now like a circle, indicating that the relationship is weak.* Accordingly, *the rounding of this ellipse also makes it difficult to determine its direction.* It's the biggest problem here. I didn't say this in the previous ones, but *when you think in this context of drawing an ellipse, I'll draw something like this, and approximately, there will be an increase-increase. But if you ask how strong relationship it has, it is weak."*

As excerpt showed, he thought drawing an ellipse surrounding all the data points to determine the trend of data points and approximate the place of the best fitting line. In addition, he interpreted (2,100) data point as behaving different from the general trend of dataset. Therefore, he preferred to exclude this data point to determine the line placement. Particularly, he explained the decision of extracting (2,100) point based on ellipse criterion. In this respect, he stated that if he included this point, ellipse drawn around whole dataset would turn into a circle and cause difficulty to identify a trend of data points. In addition, he stated that this circular shape could be interpreted as an indication of weaker association. Hence, he adjusted the line of best fit applying the ellipse criterion for seven data points (except the data point with coordinates (2,100)) and interpreted the line as displaying weak relationship between variables.

When it comes to Task 6, TE4 used this criterion as follows:

"..(draws an area) again there is a rounded shape, but the length dimension is only dimension, and this is it".

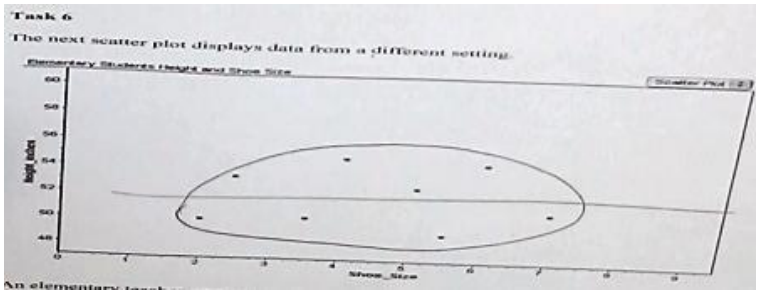

Figure 5. 4. TE4's line placement for Task6.

As data and Figure 5.4 displayed, when TE4 created an area around the scatter of data points, he found a unidirectional shape that showed variability in students' shoe sizes while students'height measures were stable. Therefore, he placed a horizontal line of best fit as dividing the area with two equal pieces. In particular, this line placement with zero-slope indicated non-association between students' height and shoe sizes.

Lastly, TE2 summarized the line placement process referring to the *ellipse criterion* for the 2nd summary question:

"We are trying to see the trend, what we call 'the trend' in dataset is the relationship between variables as the increasing-increasing, decreasing-decreasing, increasing-decreasing. While determining this trend, I am trying to see how much of this data on the linearity. *We also draw an ellipse including the data points to determine the correlation. The weaker and thinner the ellipse is, the more likely it is to show us that the line exists here. So, the logic behind drawing an ellipse is that should show me two points through which line passes (showing the points at the two endpoints of the ellipse). So as the line passes through two points, I'm actually trying to find where those two points could be. The ellipse slightly shows me their places. It is the reason I draw this ellipse. The problem starts when the ellipse starts to get rounded. When the ellipse becomes too round, you can not be sure where those endpoints are. So when the dataset is too dispersed, finding the endpoint is getting difficult. Our problem starts here. But if the ellipse is weaker, it helps me determine the endpoints."*

As excerpt indicated, he stated the first step of line placement as searching for a linear trend in dataset as increasing or decreasing. In this respect, he mentioned drawing an ellipse including all data points to determine the correlation. Then, he interpreted that the thinner the ellipse is, the more likely the best fit line is existed here. Particularly, he explained the reason of using the ellipse criterion as getting two endpoints of the ellipse that were supposed to be on the best fit line. However, when the ellipse turns into a circular shape around data, it would be difficult to identify the endpoints. In this case, the variables might have no association in between and so the line of best fit could be placed as horizontally.

In sum, each teacher educator (TE2 and TE4) used the criterion for placement of the line of best fit as drawing a shape surrounding to all data points. In particular, they thought this method as a way of handling all data points as a whole. Further, they aimed to identify the trend of the data points with use of this shape. After they determine the direction of association, they thought to place a line passing through the endpoints of the shape and so divide it symmetrically. While applying this criterion, TE2 used the term *'ellipse'* to describe the shape that he utilized. This participant referred to *'ellipse'* criterion on Task 5,

Task 6 and Task 7 including dataset with lower associated variables than previous tasks. Further, he thought the trend of data points could be easily identified for dataset having strong correlation. However, when the correlation between variables decreases, he preferred to use ellipse method to determine the trend and the best fit line. On the other hand, TE4 called this method as '*area approach'* and used this criterion for line placement for almost all task situations. In this respect, he constructed a convex area around dataset and used similar reasoning with TE2 to place the informal line of best fit. Hence, although this method does not indicate an approach supporting procedural methods, it could give an approximation for the place of the best fit line informally.

After all, the participants' dominant criteria were categorized as *the closest, sum deviation and ellipse & area approach* for this study. In particular, teacher educators who explained their line placement process using the *closest criterion* reflected the reasoning behind the least square regression method in general. Correspondingly, they stated the aim of the process to reach minimum distances between actual data values and their predicted values. On the other hand, the participants referring to the *sum deviation* provided balance of sum of deviations above and below the line. It is useful method for finding the best fit line, if this criterion applied along with the closest criterion. Lastly, creating an area or ellipse around given dataset is an efficient way of understanding the direction of association, strength of correlation and also making estimation for the place of informal line of best fit. In accordance with given information, I demonsrated dominant conceptions, their operational definitions, example excerpts, and the frequency of teacher educators who dominantly used the particular criterion in Table 5.4 below.

Table 5.4. Teacher educators' dominant criteria with their operational definitions, illustrative excerpts and frequency.

Criterion	**Definitions**		**f**
Closest TE1,	The line that is closest to all points (implies	TE1- "The method that I have applied just now is that: *the line that is the closest to the points is the right one. I just need to think about all the points, not just some.* Of course there may be some extreme values, but I try to draw *a line that passes near the majority*	7

Table 5.4. Teacher educators' dominant criteria with their operational definitions, illustrative excerpts and frequency (cont.).

Criterion	**Definitions**		**f**
TE2*TE3, TE5, TE6, TE8*, TE11	total distances from data points to line is minimized) (Casey, 2015b)	*of all these points.* I mean I will not draw the line through some of these points... the thing is: two points are enough to draw the line graph, all we do is to find two points and draw them together. But there's no such thing. These are non-linear anyway, we already realize that. Drawing a line passing through either of them is useless. *So I'm not trying to draw through any point, rather I'm trying to draw a line close to all points.* Because of that, I need to check the distance between the line and the points." (2nd summary question) TE3- "…*we are trying to minimize the sum of the squares.* The line of best fit is based on this structure. That's what they call 'the least square regression method'. *Keep the distances as less as possible, minimize the squares.* I've paid attention to this in general. What I mean by distance is, let's say the coordinates of the data point is real x or real y, the coordinates on the line are their predicted values. *I minimize the distance between them (coordinates of real and predicted x and y values), so that my model, the model that I will obtain by drawing this line segment will give the results that are close to the actual data values as possible.* x or y does not matter, $y\text{-}y_{predicted}$ is the residual for each one (data point) there. *We're trying to minimize this."* (Task1) TE8- "*It is essential that the distances from the points to the line are minimized. At the moment, I think that it will be the minimum just by looking at it***.** So I think this might be the line that represents them. Or a more special... for example, I am not affected by the context. The situation would be the same even if it was something other than a ball. So here it is important that this line cannot be horizontal, cannot be vertical, *it is essential that the vertical distances of these points are minimized.*" (Task1)	
Sum Deviation TE2*,TE7, TE8*, TE9, TE10	Deviations for points above the line to sum to the same as the sum of deviations for points below the line (they did not necessarily make measurements, but referenced this as they placed their line)(Casey, 2015b)	TE2- "There is a linear association here. Examining in a similar way, what I'm trying to calculate is that: I'm *trying to keep the residuals above and below the line approximately the same.* I suppose it would be something just like this." (Task 3) TE7- *"Rather than minimizing the distance in both, my strategy was to keep the deviation under control (in task 1)… to balance the distances above and below…* I'm always considering the deviation. In both of them." (Task 1- 2) TE8- "...I left the four points on one side, I left the four points but the numerical equation is not an essential parameter to place the line. *The important point is the balance of the sum of distances on each side***.** In fact, the way we place it now, the distances above the line were more effective, maybe. So we can move it a little more." (Task 3) TE9- "*I'm trying to center a little***.** I thought that if it passes through just one, and the error of another one is 2 and the other is 0; it would give an error of 4 units. *If it is 1, 1, 2 units will be the result. So it's reasonable to balance them. So I tried to center it approximately,* but there are three data points and here are five when it is passed from center. So I've tried to get a little bit closer to five of them, not so much, but a little bit…" (Task2	5

Table 5.4. Teacher educators' dominant criteria with their operational definitions, illustrative excerpts and frequency (cont.).

Criterion	**Definitions**		**f**
Ellipse & Area Approach TE2*,TE4	Drawing an ellipse or an area surrounding to the all data points and determining the line by dividing it symmetrically.	TE2- "We are trying to see the trend, what we call 'the trend' in dataset is the relationship between variables as the increasing-increasing, decreasing-decreasing, increasing-decreasing, I mean at first we are trying to see what is the increase or decrease amount corresponding to an increase. While determining this trend, I am trying to see how much of this data on the linearity. *We also draw an ellipse including the data points to determine the correlation. The weaker and thinner the ellipse is, the more likely it is to show us that the line exists here. So, the logic behind drawing an ellipse is that should show me two points through which line passes (showing the points at the two endpoints of the ellipse). So as the line passes through two points, I'm actually trying to find where those two points could be. The ellipse slightly shows me their places. It is the reason I draw this ellipse. The problem starts when the ellipse starts to get rounded. When the ellipse becomes too round, you can not be sure where those endpoints are. So when the dataset is too dispersed, finding the endpoint is getting difficult. Our problem starts here. But if the ellipse is weaker, it helps me determine the endpoints.*" (2nd Summary question) TE4- "I was drawing like this for this case. *I consider it as a piece of wood, and I'm trying to find a line that will divide it into two equal parts. While doing so, let me say that I could have drawn it in the concave form, but I consciously used the convex form*. In other words, from the psychological point of view, I wanted to reduce multiple data without losing their values. For example, *I imagined a piece of wood like that. While I was searching for a line that would divide this piece of wood into two, I aimed to go beyond these data points and use them without losing the values they represent.. So I created an area here, actually, we can call it as the area approach, so I divided this area into two pieces. I mean, the method as dividing the wood into two pieces is easier than considering all points at the same time"* (Task 1)	2
	*TE2 and TE8 had more than a dominant criterion for placement of informal line of best fit.		

5.2. Placing a Line for Data with Nonassociation

In this part, I aimed to explain teacher educators' approaches to placing a line for data with nonassociation. In accordance with this purpose, I focused on two data sources: 1) statements from line placement process for Task 6 regarding association between students' height and shoe sizes and Task7 regarding association between students' height and birth month and 2) line placements on given scatterplots for these tasks. Although the

line of best fit for nonassociated variables would be calculated as a horizontal line with zero-slope based on the procedural methods generally (Casey, 2015b), in this study, there was variability in slope of teacher educators' line placements for these tasks. Therefore, I specifically examined the cases of nonassociation presented on Task 6 and Task 7. In the analysis of data sources, I determined categories regarding inclination of teacher educators' line placements for Task6 and Task7. Then, I examined whether there is consistency in teacher educators' approaches to these tasks.

In the analysis of relevant data sources, I used six categories regarding inclination of teacher educators' line placements as *negative, slightly negative, zero-sloped, slightly positive, positive* and also *no placement*. In particular, four teacher educators (TE1, TE6, TE8 and TE11) placed their lines *positively* at least once, five teacher educatos (TE2, TE3, TE4, TE5 and TE 9) used *zero-sloped* line at least once, two teacher educators (TE7 and TE9) determined their lines as *slightly positive* at least once, two teacher educators (TE8 and TE10) preferred *not to place a line* for at least one task, two teacher educators (TE7 and TE11) placed their lines as *slightly negative* at least once and a teacher (TE3) determined the informal line of best fit *negatively* for once. Below, I constructed Table 5.5 representing six categories held by teacher educators, illustrative excerpts for each category and frequency of participants who placed their informal line of best fit displaying the relevant category.

Table 5.5. Categories regarding inclination of teacher educators' line placements for Task 6 and Task7, illustrative excerpts and frequency of each category.

Categories	**Illustrative Excerpts**	**f**
Positive TE1, TE6, TE8, TE11	TE1-"Now, why are we so indecisive? Has the line a positive slope, or negative slope, all because of the relationship between the variables, as well as the points seemed a little dispersed. That is why I couldn't decide. First, I looked over the scatter of data, actually. I examined the variables because the scatterplot wasn't very clear. Let's say that if there was an accumulation here, then… I would say...yes, the linear function is increasing anyway, regardless of the variables; even if the variables do not make sense to me, when I read the graph, it would already show a linear increase and I would not hesitate to say so. But now that the data seems to be dispersed, I thought over whether it is positive or negative. So I examined the variables, but I couldn't decide again. Then I went back to the graph. *I thought if I place the line just like this (positive sloped), I could make it more closer to the points. When I tried the negative sloped, the other points became more distant. If we didn't know anything about these variables, and if all we had was these*	4

Table 5.5. Categories regarding inclination of teacher educators' line placements for Task 6 and Task7, illustrative excerpts and frequency of each category (cont.).

Categories	**Illustrative Excerpts**	**f**
	points, then this would be the nearest line."(Task 7-Closest criterion) TE6- "It says primary school teacher. I mean, one will be 4th grade. How old is the child in forth grade? 10 years. *The child doesn't wear size 7 shoes when he is 10. The size he can wear is no more than 5 or 6, I think it can not be more than that.* When I examine it, I already see that the data is gathered in this area. So I believe the line should be like that. *That is, the person who has size 2 shoes is 50 cm tall, the other who has size 3.5 shoes is 51 cm tall... After all, I can say that there is a linear relationship between the shoe sizes and the students' height."* (Task6- Reflecting context criterion) TE8- *"...I draw this according to the estimated distances. I'm estimating, but I'm using the distances as well. So the line is passing through two points right now, three points left on one side, three points left on the other***.** So we can not claim that the line passes through the point, there still can be an invisible, calculable distance of that point." (Task 6- Sum deviation) TE11- "Now, of course, while drawing the line, I'm going to reason it out. *Is there a linear relationship, for example, such as his shoe size gets larger while the child gets taller? Because it shows not much of a relationship as it is dispersed. If we start to think whether it is in a positive or negative direction… I feel like, just by looking at it, it is much better (drawing a positively sloped line)…*you are doing a study here, you will create a correlation or regression line, you are examining the correlation. I believe *8 data is few*, I can ask more people to be able to make better predictions. *So I asked what could be the relationship between the height of a person and the size of the shoe; I thought as the child's height increases, the size will also get larger. Because younger kids have smaller feet."* (Task 6-Reflecting context)	
Slightly positive TE7	TE7-"It is too dispersed. Oh, *I thought to draw like this (she placed horizontal line). It can be. But, I know from daily life that there is a relationship. But, when we draw like this (horizontal) and then make a little inclination, again it can be. But, it is not significant."* (Task6- Sum deviation and reflecting context)	2
Zero-sloped TE2, TE3, TE4, TE5, TE9	TE2-" So when I tried to draw an ellipse, *I saw that the ellipse I drew indicated very weak relationship…*We don't have difficulty in the others because the trend is very clear. But here, *the trend is not very clear. So, we think of an ellipse like that, and when we also think of the ellipse's endpoints, I thought this would fit better....*"(Task 6- ellipse criterion) TE2- "*The condition of having zero-slope is equivalent to the condition of having no association*. Normally, we do not directly associate the slope and relationship, however, *having a zero slope here is equivalent condition with having no association.*" (Task7-ellipse criterion- thought of horizontal line of best fit for nonassociation) TE3-"*I have constructed a weak, almost zero, and positive relationship.* Then, I constructed my line of best fit as *approximately 4 points above and 4 of them below,* but more importantly, I created *a line in order to minimize the distances between the line and those data points."* (Task6- closest, equal number) TE4-"..(draws an area) again *there is a rounded shape, but the length dimension is only dimension, and this is it"* (Task6- area approach) TE5- "*There's a very weak relationship. Because the case of the nonassociation requires to have a horizontal curve. So as r to be zero. If r is zero, then yes, it requires to be a horizontal line***.**" (Task 6- thought of horizontal line of best fit for nonassociation) TE9- "Now, *we do not expect any relationship between the birth month and the height***.** Is there any relationship? Maybe there is.. I mean, tall people may wear large sized shoes, but *I don't even examine the points in that case, I may immediately think of zero… I balance it again, but I have to say that I'm more prejudiced than ever before.. but I wouldn't think so if the points were the same as the previous ones. I tried to center the line again.* I wanted to make it suitable for the possibility of giving errors. Because if it is moved even a little, the possibility of errors would be increase, and maybe there will be no relationship	5

Table 5.5. Categories regarding inclination of teacher educators' line placements for Task 6 and Task7, illustrative excerpts and frequency of each category (cont.).

Categories	**Illustrative Excerpts**	**f**
	again…*I just tried to balance, but like I said before, I tried to place it in a parallel way*." (Task 7- reflecting context- sum deviation)	
Slightly negative TE7, TE11	TE7-*"…there is no very serious, strong negative relationship. I mean, you know , looking at such a data, it does not give me that negative feeling as previous ones. With the benefit of context, i can make a little more decision***."** (Task7- Sum deviation and reflecting context) TE11- *"..this context does not make sense to me that would offer a relationship***.** Therefore, I will choose such a line for data points that will have equal distances to the line. It seems as if it is good by just looking at it, lets paste it. *In the context, I do not think there is a relationship, but, when I look at the dataset, I might draw something like that. As I said again, when we evaluate the distances to the line."* (Task7- Closest)	2
Negative TE3	TE3- *"We might exclude first data point. If we discard this data, I put something like this (negatively sloped).* I think we can discard it here ... when we look at the context, how useful it is to include this point, if our goal is to get benefit from it. So if our goal is to construct a relationship between the height of the students and the birth month by using statistics, *this point is somewhat misleading*. The line of best fit that we will create using this point and the regression equation will be misleading. At least, *when we discard this point, we get closer values to the actual data by using 7 data points.* But as I said, there is no definite answer here." (Task7- extracting a data point with consideration of outlier- Closest)	1
No placement TE8, TE10	TE8- "*...it is not possible to search for such a relationship here. If we try to fit a line, mathematics could provide it***.** What gives? It gives something having 0.05 correlation coefficient value. *But, statistics or interpretations could not be used based on the context. The best thing to do here is to decide that regression could not be searched, searching for a regression equation is not correct after seeing the context and the data."* (Task7) TE10- "*Don't we expect that who is the tallest would have the larger size of shoe? In general, so it is. But, here, it will be the opposite*. Okay, now the shoe size increases as the height increases, right?. Isn't it so**?** *Let's see if it's increasing. No, it's weird.* Wait a minute… **I***t doesn't match with what we expected, so what do you think?* It is expected that the tallest person has to have larger sized shoes, if there are no abnormalities. Here, who is 55 feet tall has smaller sized shoe, that is uninterpretable**.** *So if you draw a linear.. yes, you can draw a line, but the result can not be interpreted. I don't know if someone can explain this relationship. It would be an abnormal situation. In other words, naturally, we cannot support this relationship theoretically.* In other words, the data can be obtained, but an erroneous data may be obtained here…" (Task 6)	2

As Table 5.5 demonstrated, teacher educators reflected different categories of inclinations while determining the best fitting lines on Task 6 and Task 7 having nonassociated data. In particular, they referred to different criteria and methods while explaining their line placement process. In order to understand their reasoning behind their line placements for these tasks of nonassociation better, I focused on the categories teacher educators consistently used while placing their lines on each task situation. In particular, I

constructed Table 5.6 showing teacher educators' categories for inclination of their line placements for Task 6 and Task7 below.

Table 5.6. Categories for inclination of teacher educators' line placements for Task 6-7.

	Categories for Task6	Categories for Task7
TE1	*Positive*	*Positive*
TE2	*Zero-slope*	*Zero slope*
TE3	Zero-slope	Negative
TE4	*Zero- slope*	*Zero slope*
TE5	*Zero- slope*	*Zero slope*
TE6	*Positive*	*Positive*
TE7	*Slightly Positive*	*Slightly negative*
TE8	Positive	No placement
TE9	*Zero slope*	*Zero Slope*
TE10	*No placement*	*No placement*
TE11	Positive	Slightly negative

As the Table 5.6 showed, two teacher educators (TE1 and TE6) consistently placed their lines *positively*, four teacher educators (TE2, TE4, TE5 and TE9) consistently determined their lines with *zero-slope*, a teacher educators (TE7) consistently made *slightly sloped* line placements and a teacher educators (TE10) rejected to determine a line for each task with nonassociated variables. In the following paragraphs, I give more detailed information about their reasoning for their perspectives.

Teacher educators who consistently placed their lines *positively* (TE1 and TE6) referred different criteria and methods for Task 6 and Task 7. In particular, for Task 6, TE1 stated that:

"This is *dispersed dataset, so it won't come out linear. I mean, if we look at the correlation, it seems to be closer to zero. When I make a horizontal line, they actually look like having equal distances, but then there are a few points like this, at least that's very close to the line. I think it is better (positively sloped line). If we look at the context, it would be expected that a taller person would have a bit longer foot, right? I mean, it's a more positive relationship ... what context says not always consistent with what the data says.* For sure there are 8 people here. I mean, it doesn't have to be, but this affects a person's decision"

As the data suggested, TE1 firstly interpreted dataset as reflecting no correlation between variables and thought placing the best fit line horizontally. However, while examining the data points, she also thought that when she placed the line with positive slope, some data points seemed to be closer. Therefore, she determined the best fit line on Task 6 with positive slope referring to the closest criterion for placement. After that, she also supported her line placement with use of context information. For her, *"...it would be expected that a taller person would have longer foot…"* Therefore, she concluded that the best fit line would reflect positive relationship in accordance with her expectations regarding the contextual information. However, she stated that 'what the context says was not consistent with what the data shows' for this task. In particular, she further explained this inconsistency with the small number of data points given on scatterplot. Hence, although she referred to the closest criterion while placing the line, contextual information also affected her decision of line placement.

Similarly, TE6 also made interpretation focusing more on contextual information while placing the best fit line on Task6:

"It says primary school teacher... It means, they will be up to 4th grade in primary school. How old is a 4th grade child? S/he will be 10. *They do not wear shoe size as 7 when they are 10 years old. The shoe sizes s/he can wear no more than 5 or 6, so I think it can not be more than that.* When I examine it, I already see that the data is also accumulated here (around smaller shoe sizes). I think, it should be such a line. *That is, the*

person who has shoe size as 2 is 50 cm tall, the other who has shoe size as 3.5 is 51 cm tall... After all, I can say that there is a linear relationship between the shoe sizes and the students' height."

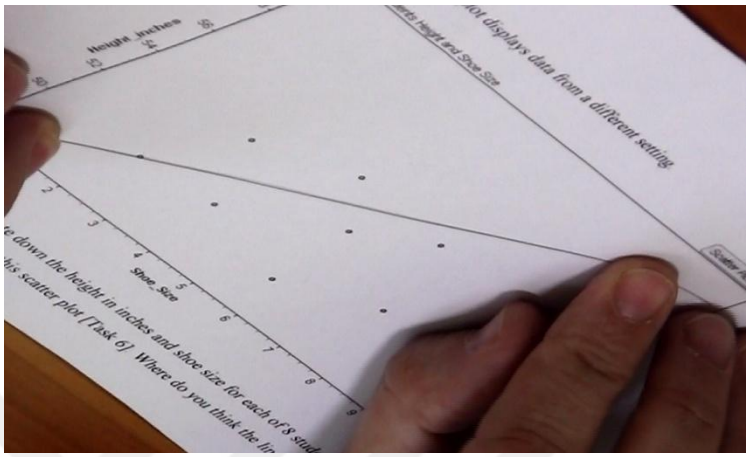

Figure 5. 5. TE6's line placement for Task6.

As the data and Figure 5.5 indicated, she used contextual information to disregard some data points indicating higher values of shoe sizes. Particularly, she determined her line placement with positive slope focusing on the subset of data points having smaller values for shoe size variable. Therefore, her criterion for line placement on Task 6 was reflecting context dominantly.

On the other hand, these participants was also determined positively slope lines for Task 7. However, in this case, they explained their reasoning for line placement focusing more on given dataset rather than contextual information. In particular, TE1 stated that:

"Now, why are we so indecisive? Has the line a positive slope, or negative slope, all because of the relationship between the variables, as well as the points seemed a little dispersed. That is why I couldn't decide. First, I looked over the scatter of data, actually. I examined the variables because the scatterplot wasn't very clear. Let's say that if there was an accumulation here, then… I would say...yes, the linear function is increasing anyway, regardless of the variables; even if the variables do not make sense to me, when I read the graph, it would already show a linear increase and I would not hesitate to say so. But now that the data seems to be dispersed, I thought over whether it is positive or negative. So I examined the variables, but I couldn't decide again. Then I went back to the graph. *I*

thought if I place the line just like this (positive sloped), I could make it more closer to the points. When I tried the negative sloped, the other points became more distant. If we didn't know anything about these variables, and if all we had was these points, then this would be the nearest line."

As the data showed, she thought that the dataset was given as dispersed to identify a linear direction positively or negatively. In addition, when she considered the relationship between context variables, she also could not find an evidence for direction of association. Therefore, she was in two minds about whether the line should be determined as positively or negatively sloped. Finally, she concluded that the line reflecting positive direction of association would be closer to the data points compared to the line with negative direction. Therefore, she determined the best fit line on this task referring to the closest criterion.

Similarly, TE6 also mentioned:

"So *we can say that it is the one which is the closest to the line...* or the line which *passing through the area closest to those points."*

As the excerpt indicated, she referred to the closest criterion for placing a line on Task7. Hence, similar to TE1, she also thought that she could determine the closest line to the data points when she placed it positively.

In sum, teacher educators placing their lines positively on each task with nonassociation used different reasoning for placement. In particular, for Task 6, they utilized contextual information while giving decision of direction of association. Therefore, they used *flexible approach* in line placement based on their expectations for the relationship between variables. On the other hand, for Task 7, they aimed to find the closest line to the data points. In this case, although they evaluated the relationship between the context variables to determine the best fit line, they could not give decision since the variables were given as unrelated with each other. Therefore, they referred to the deterministic perspective that supports the idea behind procedural methods on Task 7.

In addition, the teacher educator (TE7) who made line placements as *slightly sloped* for Task6 and Task7 explained her reasoning in the following.

For Task 6, she stated that:

It is too dispersed. Oh, *I thought to draw like this (she placed horizontal line).* It can be. But, *I know from daily life that there is a relationship. But, when we draw like this (horizontal) and then make a little inclination, again it can be. But, it is not significant."*

As the data indicated, she thought to place horizontal line of best fit at first, since the data is given as so dispersed. However, when she took into consideration of contextual information, she expected to have positive association between variables. Therefore, she preferred to make slightly sloped line placement. However, she also pointed that the association between variables displayed with this line was not significant.

Similarly, TE7 mentioned on Task 7 such that:

"*…there is no very serious, strong negative relationship*. I mean, you know , looking at such a data, *it does not give me that negative feeling as previous ones. With the benefit of context, i can make a little more decision."*

As data showed, she made similar reasoning with the previous task on Task 7. Again, she thought that there is no significant association between variables. However, she placed the line as slightly slope in the light of the idea derived from contextual information.

To sum up, TE7 considered placing almost horizontal line of best fit for each task given with nonassociation. Through she aimed to balance deviations above and below the line, she also evaluated the relationship between context variables in the line placement process. In fact, her line placement indicated that she has the knowledge of using

horizontal line of best fit with almost zero-slope for nonassociated variables. Therefore, her approach was in accordance with statistical procedures.

On the other hand, the teacher educator (TE10) who passed Task 6 and Task 7 with *no placement* interpreted these cases as it is explained below.

For Task 6, he stated that:

"*Don't we expect that who is the tallest would have the larger size of shoe? In general, so it is. But, here, it will be the opposite.* Okay, now the shoe size increases as the height increases, right?. Isn't it so**?** *Let's see if it's increasing. No, it's weird.* Wait a minute… *It doesn't match with what we expected, so what do you think?* It is expected that the tallest person has to have larger sized shoes, if there are no abnormalities. Here, who is 55 feet tall has smaller sized shoe, that is uninterpretable**.** *So if you draw a linear.. yes, you can draw a line, but the result can not be interpreted. I don't know if someone can explain this relationship.* It would be an abnormal situation. In other words, naturally, *we cannot support this relationship theoretically.* In other words, the data can be obtained, but an erroneous data may be obtained here…"

As the excerpt suggested, he expected that there would be positive association between given context variables. However, he identified that data points did not scattered as if there exists an association. Therefore, he thought that there was inconsistency between data points and given contextual information. Hence, he refused to place a line for this task.

In particular, TE10 elaborated on Task 7 as:

"*Unrelated two variables.* We think that…D**oes** *this variable have an effect on this? There could be earlier assumptions, studies, a preliminary data, considering any of this you can say whether it is a variable*… Sometimes, okay, there may not be strong relationships, but the situation may be the one which can be explained. *I mean, in this variable there is no such thing as the date of birth, I mean, people born on that date will be…. There is not any logical explanation of this."*

Data pointed that, he thought that the variables did not make sense to search for an association. In this respect, he stated that there is no logical explanation of the association between the variables as students' birth month and their height. Therefore, he refused to place a line on this task, too.

In sum, TE10 consistently made decision of not to place an informal line of best fit for the tasks with nonassociation. However, he explained his reasoning on each task differently. In particular, while he was referring to the inconsistency between the relationship inferred from contextual information and scatter of data points on Task6, he evaluated the variables as unreasonable to search for an association on Task 7. Therefore, he reflected his ideas about contextual information dominantly while reasoning on these tasks. Therefore, he used flexible approach focusing more on contextual information while giving decision rather than deterministic perspective focusing on procedural methods.

Lastly, four teacher educators (TE2, TE4, TE5 and TE9) consistently placed *zero-sloped* lines for Task 6 and Task 7 with nonassociation. In particular, for Task 6, TE4 stated that:

"..(draws an area) again *there is a rounded shape, but the length dimension is only dimension,* and this is it."

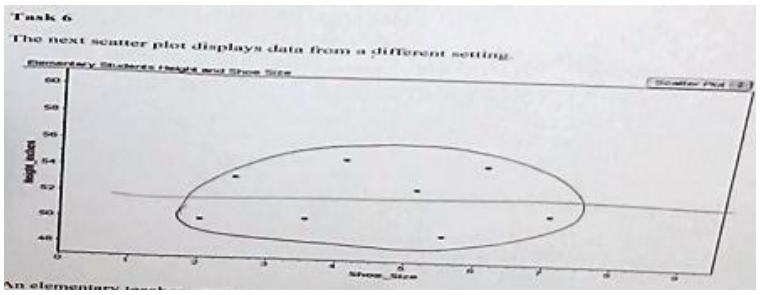

Figure 5. 6. TE4's line placement for Task 6

As the data and the Figure 5.6 indicated, he determined the informal line of best fit on Task 6 with zero-slope. In particular, he explained his process focusing on his dominant criterion for placement as area approach. However, he did not make any interpretation relevant to the association between variables.

In particular, TE5 described the dispersed dataset on Task6 as an indication of nonassociation. Then, he stated that:

"There's a *very weak relationship*. Because *the case of the nonassociation requires to have a horizontal curve. So as r to be zero. If r is zero, then yes, it requires to be a horizontal line."*

Data pointed that, he thought that the best fitting line should be horizontal since there is no association between variables. Also, he supported this idea with the information such that if the correlation coefficient (r) is zero, then the slope of the best fit line would be calculated as zero.

Lastly, TE9 also interpreted that:

"I mean, since we firstly thought that *it is not increasing*, since I thought *that there is no relationship, I placed it as passing through the center.. but I made it close to zero..."*

As data indicated, he made interpretation of nonassociation focusing on the scatter of dataset. Then, he applied his dominant criterion, sum deviation, while placing the informal line of best fit. Particularly, he described his line placement on this task as passing through the center of data with almost zero-slope.

On the other hand, TE2 elaborated on Task 7 such that:

"*The condition of having zero-slope is equivalent to the condition of having no association.* Normally, we do not directly associate the slope and relationship, however, *having a zero slope here is equivalent condition with having no association."*

As data suggested, he interpreted that there is a connection between zero-slope line of best fit and the dataset having no association.This explanation supported the idea behind the statistical hypothesis test for slope of regression line that used to make inference whether there exists an association or not.

Also, for Task 7, TE5 mentioned that:

"…If I comment on it (referring to the line of best fit he placed) in this context, it says that *whatever the date of birth is, the height of a person is stable.* When I look at it now, *it would give us the average, the average height, it's actually possible."*

As the data suggested, he interpreted the horizontal line of best fit that he placed for Task7 as representing the average of students' height. In particular, since the dataset is given as nonassociated, then the line of best fit would provide similar estimations with taking an average of interest variable.

Finally, TE9 also placed horizontal line of best fit on Task 7. Then, he explained his reasoning as follows:

"Now, *we do not expect any relationship between the birth month and the height.* Is there any relationship? Maybe there is.. I mean, tall people may wear large sized shoes, but *I don't even examine the points in that case, I may immediately think of zero… I balance it again, but I have to say that I'm more prejudiced than ever before.*. but I wouldn't think so if the points were the same as the previous ones. I *tried to center the line again.* I wanted to make it suitable for the possibility of giving errors. Because if it is moved even a little, the possibility of errors would be increase, and maybe there will be no relationship again…*I just tried to balance, but like I said before, I tried to place it horizontally."*

As the data indicated, TE9 utilized the contextual information to make interpretation of nonassociation on Task7. In particular, he stated that he immediately think of zero-sloped line even before examining the data points in this case. After he decided to place he line horizontally, he also examine the data points to provide balance of residuals. Hence,

he used sum deviation criterion along with reflecting context criterion while placing horizontal line of best fit for data with nonassociation.

In sum, the teacher educators placing horizontal line of best fit on both tasks (TE2, TE4, TE5 and TE9) demonstrated that the dataset shows no association between variables. In addition to this, for Task 6-7, TE2 used ellipse criterion, TE4 used area approach and TE9 used sum deviation criterion for placement that they also applied on previous task situations dominantly. Further, TE2 and TE9 associated having zero-slope line of best fit with the data displaying nonassociation. However, TE4 obtained horizontal line of best fit with the benefit of his dominant criterion and did not made interpretation for association between variables. On the other hand, TE5 made interpretation of nonassociation when he examined the given dataset. Then, he placed the informal line of best fit horizontally to indicate no association between variables. Also, he used the relationship between having zero value of correlation coefficient (r=0) and the slope of regression line to justify his line placement. Further, he interpreted the horizontal line of best fit as showing average value of interest variable whatever the other variable value is. For him, the resulting line would give the same estimations with taking average of interest variable using univariate data.

After all, Task 6 displaying association between students'height and shoe sizes and Task 7 showing association between students' birth month and height are included in the interview protocol to understand teacher educatos' perspectives in case of nonassociation. In particular, participants are expected to place horizontal line of best fit not only to describe the present data, but also to make statistical inference about having association or nonassociation between variables. In particular, while four teacher educators (TE2, TE4, TE5 and TE9) consistently made placement of horizontal line of best fit, three of them (TE2, TE5 and TE9) made inference of having no association between variables on these tasks based on given dataset and their line placements. In addition, TE7 who placed slightly sloped lines of best fit on Task 6-7 interpreted that there is no significant association between variables. In particular, she also made interpretation beyond describing the dataset.

In addition, these tasks also gave information of teacher educators' perspectives as deterministic (focusing on procedural methods) or flexible (focusing on the contextual information) while placing the best fitting line in case of nonassociation. In particular, the results showed that some participants (TE1, TE6, TE8 and TE11) reflected flexible approach, since they dominantly used contextual information while placing the line. In addition, the only participant (TE10) who refused to make line placement on these tasks also explained his reasoning referring to the contextual information. On the contrary, some participants (TE2, TE3, TE4, TE5, TE7 and TE9) dominantly used deterministic perspective as referring to their dominant criterion for placement while placing their lines on Task 6 and Task 7. In fact, teacher educators who dominantly used deterministic perspective as focusing on given scatter of data points and applying the criterion behind procedural methods made better line placements for different task situations. Additionally, four teacher educators (TE2, TE5, TE7 and TE9) evaluated almost zero-slope line of best fit to make interpretation of association between context variables. In this respect, they used flexiable approach while making interpretation of slope of the regression line as recommended in the previous studies.

5.4. Slope Conceptualization and Line of Best Fit

Slope concept is studied within linear functions before introduction of linear regression. Therefore, some questions are arisen whether slope conceptualizations have an effect on the conceptualization of the linear regression. Particularly, examining the answers to the 9th and the 10th questions, I looked for teacher educators' reasoning while finding the estimate value of slope of their line placements on Task 1 and Task 3. Further, I also asked for interpretation of meaning of slope value in given context within these tasks. In addition to the relevant information depicted from the 9th and 10th questions, I also examined statements from line placement process on different task situations whether they interpreted the association using slope meaning for the line of best fit. In the analysis of these data, I firstly determined slope conception categories teacher educators referred throughout the interview process. Then, I identified dominant slope conceptions according to the frequency of their reference to these categories.

The analysis of these data sources addressed the research question "What slope conceptualizations do teacher educators use for conceptualization of informal line of best fit?" In this respect, I found evidences for ten slope conceptualization categories, namely, *geometric ratio, algebraic ratio, physical property, functional property, parametric coefficient, trigonometric conception, real world situation, determining property, behaviour indicator and linear constant*. In the following, I constructed Table 5.7 displaying teacher educators' use of slope conceptualization categories along with the dominant conceptions for the best fit line. Particularly, I also depicted the dominant slope conceptualizations among teacher educators having the same dominant conceptions. By doing this, I looked at the frequency of teacher educators referring to the relevant slope conceptualizations who have the same dominant conceptions for the best fit line.

Table 5.7. Teacher educators' slope conceptualizations along with their dominant conceptions of informal line of best fit.

TE	Geometric Ratio	Algebraic Ratio	Physical Property	Functional Property (FP)	Parametric Coefficient (PC)	Trigonometric Conception	Real World Situation (RW)	Determining Property	Behaviour Indicator (BI)	Linear Constant	Dominant slope conception	Dominant conception for line of best fit
TE1		X	X	X	X		X		X	X	BI	Representer
TE4	X		X	X	X		X		X	X	BI	Representer
TE6		X					X		X	X	BI	Representer
TE7			X	X	X	X	X		X	X	BI	Representer
TE8					X	X	X	X	X	X	RW	Representer
TE2	X			X			X		X		RW	Predictor
TE3		X		X			X		X		BI	Predictor
TE5		X	X	X	X		X		X		RW&BI	Predictor
TE11	X				X		X		X		BI	Predictor
TE9		X		X	X		X		X		RW	Signal
TE10		X		X	X		X		X		FP& PC	Signal
f	3	6	4	8	8	2	11	1	11	5		

In addition, I examined the data whether there was an association between participants' use of slope conceptualizations and their conceptions of informal line of best fit. As the Table 5.7 showed, there are some commonalities of slope conceptualizations for teacher educators having the same dominant conceptions. In particular, teacher educators having *representer* conception (TE1, TE4, TE6, TE7, TE8) for informal line of best fit dominantly used the slope conceptions as *linear constant* and *parametric coefficient* along with *real world situation* and *behaviour indicator* conceptions; teacher educators having the *predictor* conception (TE2, TE3, TE5 and TE11) for the best fit line dominantly referred to the slope conceptions as *functional property* along with the *real world situation* and *behaviour indicator*; and also both teacher educators having the *signal* conception (TE9 and TE10) dominantly referred to the slope conceptions as *algebraic ratio, functional property, parametric coefficient, real world situation* and *behaviour indicator*.

Particularly, as the Table 5.7 suggested, each teacher educator not only made interpretations relevant to the *behaviour indicator* and *real world situation* conceptualizations of slope generally but also almost all of them had either of these conceptions as their dominant slope conceptions. In the following paragraphs, I give some illustrative excerpts and relevant explanations from teacher educators who referred to these slope conceptions in their interview process.

For 9th question asking for interpretation of slope of the best fit line on Task1 that displays relationship between golf ball's bounce and drop height, TE1 referred to the *behaviour indicator* along with *real world situation* conceptions as follows:

"*What does the slope represent? Rather than being positive for example...* if it was x = y, its slope would be 1. *This slope is smaller than one. That makes sense, too, because the ball will not bounce as high as the height I drop from, it will bounce less*. So, my slope, I mean... x values need to be higher than y. Thus, I need to obtain a line which is below the x=y line. I mean, I need to obtain a line that have lower slope."

As the excerpt suggested, she firstly stated that the slope value indicates positive association based on the *behaviour indicator* conception. In addition, since the slope value was calculated as 27/38 lower than 1, she interpreted that the golf ball would not bounce up to its drop height. In this respect, since he interpreted the slope meaning in given

context, he also used the *real world situation conception*. Hence, she concluded that the line of best fit for this association would have lower slant than the equation x=y.

For 10th question asking for the slope meaning for the best fit line on Task 3 that shows association between number of audiences and ticket prices, TE8 made interpretation such that:

"So *there is a decrease, but this decrease is not... not equal.* It is not a high correlation, but it can be said that there is a strictly reasonable decrease. *Filmmakers can benefit from this. They can benefit while increasing the prices..."*

As the data suggested, he started interpretation with description of decreasing trend of the dataset based on *behaviour indicator conception* of slope. In addition, he stated that decrease in a variable with increase of other variable did not have a constant rate of change. Particularly, he appreciated the variation in the amount of change in variable values referring to the correlation between variables. Therefore, he used *flexible approach* while interpreting the slope meaning with consideration of approximate rate of change between context variables. Also, he made interpretations for the slope value referring to the contextual information. Therefore, he also evaluated the real world situation.

In addition, TE5 used the behaviour indicator and real world situation conception to make comparison between Task3 and Task 4. In particular, he stated that:

"*What does this line show us?* Again, we see that *as the ticket price increases, the number of audiences is decreasing.* There is a *slightly more dramatic decrease comparing to the other graph (Task 3).* The ticket price was very cheap in Task3 at the beginning. The number of audience may have been less, that is to say the ticket price was very cheap. In Task 4, for example, the initial ticket price is not too cheap, when compared to the other. *So the decrease in Task 4 was much more than the other. When the ticket price increased, the number of audience decreased more."*

As the data pointed, TE5 mentioned the increasing-decreasing trend of the line based on the *behaviour indicator* conception. Also, since he used contextual information to explain the dynamic situation between ticket prices and number of audiences, *real world situation* conception was also used. In particular, he described the line of best fit for Task 4 as showing more dramatic decrease in number of audiences with increase of the ticket prices compared to Task 3.

Particularly, all participants interpreted the slope value as representing increasing, decreasing or horizontal trend of the best fit line based on the *behaviour indicator* conception. In addition, each teacher educator also elaborated on the slope value regarding the association between context variables in physical or functional situation referring to the *real world situation* conception. In this respect, the results showed that teacher educators had a tendency to interpret the slope value of the best fit line in given contextual information as it is recommended for statistical settings.

In addition to behaviour indicator and real world situation conceptions, all teacher educators (TE1, TE4, TE6, TE7, TE8) having *representer* conception for informal line of best fit dominantly used the slope conceptions as *linear constant* and *parametric coefficient*. In order to elaborate on *the linear constant conceptualization of slope***,** I share statements from line placement process for different task situations. In this respect, TE1 made interpretation on Task 1 such that:

*"Some points seem as if they were on the line, I mean the first, third, fifth, sixth and seventh points. So they seem linear***.** Those, for example (second data point), are a bit below the line which I've pictured. Other than that, others seem to be closer. *Why have I pictured such a line? It's like those six points seem as if they are on a line. Now if this were exactly linear, all the points would be on that line. So I thought that as many points as possible should be on or near this line."*

As the excerpt demonstrated, she evaluated given data points on the scatterplot with expectation of perfect linearity. Therefore, she mentioned that some data points were situated as linear and others deviating from this direction. This interpretation referred to the property of slope in linear functions as a constant value for each point based on the l*inear constant conceptualization* of slope. While this perspective was creating a hindrance for some students in previous studies, TE1 reflected it to the line placement process through using *closest criterion* that supports the reasoning behind the procedural methods to find the regression equation.

In particular, TE4 used the linear constant conception of slope on Task 4 as follows:

"I'm trying to pass from the majority to minority without losing any data. Dealing with eight data is more difficult than dealing with two data. So I'm trying to divide it into two parts. I know what I want to draw is line. I want to create a line. *What is line? It is*

straight. What are the characteristics of a line? It passes through two points. If I can find two points, I can draw that line through them in the most optimal way."

As data suggested, he described the line as a straight figure passing through at least two points. In this respect, he aimed to reduce the dataset including eight data points varying to some extent into two points. By doing this, he thought to place the line of best fit as passing through these two points. In particular, on this task, he reflected *linear constant conceptualization* based on which the slope was defined as a unique value for straight figures.

Also, TE7 also referred to the slope value as linear constant on Task4:

"…now, if I draw the line like this (showing the closest line to the points at the beginning and at the end), since there will be too much deviation, the number of points is not important ... *Look, if those didn't exist (4-5-6.points), these would be linear (1-2-3-7-8. data points),* at least close to linear..."

As data showed, she mentioned that data points having minimum (1,2,3. data points) and maximum x values (7,8. data points) was close to be linear. In particular, this interpretation indicated that the slope value calculated with the use of these data points would give similar values. Hence, since the slope was defined as the constant property for points on the straight figures based on *linear constant conceptualization*, I thought that this explanation led to the use of this conception while reasoning on the placement of informal line of best fit.

Lastly, TE8 referred this slope conception while evaluating the students' line placements on Task 8. In particular, he metioned:

"There is something interesting here, this (Barbara), for example, chose to pass it through the points. *As many points as possible.. At first, I thought this way was better, but I've changed my mind, because its distance from the points on the right ... so it is parallel to it (a little above), if it were such line that would be better.* Apart from choosing the right slope, the slope within the regression equation of y= mx + n, the value of n is too weak. n is the point where the line crosses the axis."

As data suggested, he firstly evaluated Barbara's line as a better placement than other students'. In this case, he gave his decision based on the comparison of slope values for students' line placements. In particular, although he thought that the slope of Barbara's line

was appropriate to show the association between variables, he did not feel confident for the predicted values of y variable on this line. Also, he identified that some data points had higher vertical distances from Barbara's line placement. Therefore, he thought that a line that is parallel to the Barbara's placement might be better. In this case, he thought the slope as "*the property that is unaffected by translation*" based on the *linear constant conceptualiztaion* of slope.

According to the *linear constant conception*, slope is defined as a constant property that is relevant to straight figures and unaffected by translation. In this respect, four teacher educators (TE1, TE4, TE6 and TE7) having *representer conception* used the *linear constant* conception of slope at least once as focusing on data points that seem collinear on given scatterplot while placing the best fitting line. On the other hand, TE8 also used slope concept referring to the linear constant property that is unaffected by translation while examining the students' line placements. In particular, he suggested that even if the Barbara's line placement was passing through three data points, there were some points that were distant from the line. Therefore, he suggested to replacing the line of best fit as parallel with the initial line placement using the linear constant property of slope.

On the other hand, four teacher educators (TE1, TE4, TE7 and TE8) having representer conception mentioned slope value as a ***parametric coefficient*** on the linear regression equation. As an example, TE4 interpreted slope of Task 6 and Task 7 as follows:

"In both of them (Task 6 and Task 7), *y=mx+p, m is near to zero*."

For Task 6-7, he used the area approach criterion to obtain the best fit line for data with nonassociation. In particular, he constructed an area surrounding whole dataset and obtained horizontal line of best fit for these tasks. In this respect, he represented the slope value (m) of his line placement as the parametric coefficient of the linear equation y=mx+p as close to zero value.

In addition, TE8 also made interpretation on 9th question referring to parametric coefficient:

"That is, about this slope... it's approximately between 40 and 45 degrees. *I probably said, something like, 0.8 or so.* Of course, my analytical skills are also effective here. For example as I *see it, y = 0.8x + 7. For X = 0, No negative (referring to the constant), but here it (the line) will cross this (y-axis) below."*

As the excerpts showed, after he determined the slope of best fit line on Task1, he made interpretation reflecting her analytic perspective by using this slope value as a *parametric coefficient* of linear equation.

Particularly, *parametric coefficient* was also one of the most frequently seen slope conceptions for teacher educators having representer conception. They used this conception by substituting the slope value on a linear equation. In fact, the parametric coefficient conception thought as reflecting analytic perspective of slope. Particularly, although the participants with representer conception were focusing more on visual representation of whole dataset, they also utilized the slope value as a coefficient on regression equation for different tasks and questions.

In sum, teacher educators having representer conception dominantly referred to slope of the line of best fit as *parametric coefficient* and *linear constant* along with *behaviour indicator, real world situation conception* at most. Although teacher educators having different conceptions for the best fit line mentioned writing an equation of linear regression with reference to the slope value as the *parametric coefficient* generally, dominancy of *linear constant conception* was arised only for the participants having *representer* conception. Therefore, there might be an association between having *representer* conception and the use of *linear constant conceptualization* along with *behaviour indicator* and *real world situation conception* of slope.

On the other hand, participants having the *predictor* conception (TE2, TE3, TE5 and TE11) for the best fit line referred to the slope meaning as *functional property* along with the *real world situation* and *behaviour indicator* at most. In particular, in the following, I share statements from these teacher educators in order to understand characteristics of *functional property* conception of slope better.

For 9th question, TE3 stated that:

It says that: *the bounce height will increase by 0.7 cm in response to each unit increased in the drop height. We predict that the ball which is dropped every 1 cm higher will bounce by approximately 0.7 cm higher ... The value of the slope we found gave the estimated gradient change in y for each unit changed in x...”*

As the excerpt showed, he explained the slope value in context as the rate of change between golf ball drop height and bounce height regarding *functional property* of slope. In particular, he stated that a unit change in drop height would be associated with approximately 0.7 unit increase in bounce height with reference to the *real world situation conception* of slope. Also, he mentioned slope of the informal line of best fit as *predicted rate of change between given variables*. In fact, since the dataset includes inherent variation, considering the slope value as an approximation is recommended while studying in statistical setting. Particularly, TE3 reflected this flexible approach rather than deterministic view that defines the slope as *the constant rate of change between variables* in mathematical setting.

Similarly, for 10th question, TE2 stated that:

“We can understand from the context that; I'm going backwards this time, a *decrease of 0.4 units within the number of audience corresponds to a 1 unit increase within this.* I mean in here. *The equivalent of a $ 1 increase.. in our calculation, if approximately 0.4 billion audience stop watching, it corresponds to a $ 1 increase in ticket price. It shows how much decrease equals to how much increase.”*

As the excerpts suggested, he interpreted the slope value for this question as indicating change in a variable with unit change of other variable based on the *functional property conception* of slope. In addition, he considered contextual information to make interpretation for this dynamic situation on given task condition. Therefore, I interpreted that he used the *functional property* in conjunction with *real world situation*. Additionally, in parallel with TE3, TE2 also mentioned the slope value calculated on this task as an approximation to the actual change in variable values. Therefore, I concluded that he also used flexible approach while interpreting the meaning of slope in context.

Lastly, TE5 commented on 10th question such that:

"Suppose it is -2/80. That will be -0.025. *When there is an increase of one unit in ticket price, number of audience* ... Pardon me, we need to multiply the number of audience by one million. *So for each dollar the ticket price is increased by, the number of audience is decreased by 2500... Now, this is an estimation, but the actual data does not show this of course, it is not very regular***.** The level of negative relationship is also very weak."

As data suggested, he stated that the calculated slope value showed the change in a variable for each unit of increase in other variable. In other words, he interpreted the slope as indicating the covariation between two variables regarding *functional property conception.* In addition, since he interpreted these functional situations referring to the contextual information, *real world situation* conception was also depicted from his statements. Also, he stated it as *average and predicted* rate of change between variables on 10th question. Particularly, he explained this interpretation based on the inherent variation in dataset. In fact, since there are deviations from the line of best fit, interpretation of association based on the slope value would be as *approximate* or *predicted* rate of change between given variables in statistical setting.

In sum, teacher educators having predictor conception dominantly used slope as *functional property* along with *behaviour indicator* and *real world situation* conception at most. Particularly, the teacher educators who interpreted the association using *functional property* and *real world situation* conceptions of slope simultaneously focused on rate of change between context variables. In fact, using these conceptions of slope in tandem could enhance understanding of association between contextual variables. In addition, these participants considered slope value of the line of best fit as a *predicted or approximate* value for given dataset varying to some extent. Particularly, since the rate of change between variables is not constant because of inherent variation, using this flexible approach while interpreting the slope of line of best fit in given context is suggested for statistical setting.

Lastly, both teacher educators (TE9, TE10) having the *signal* conception referred to the slope conceptions categories as *algebraic ratio, functional property and parametric coefficient* along with *real world situation* and *behaviour indicator*. In the following paragraphs, I give illustrative excerpts to indicate their slope interpretations while working on the informal line of best fit.

Each teacher educator having signal conception used *algebraic ratio* while calculating the estimated slope values for 9th and 10th questions. In particular, for 9th question, TE9 stated that:

"(Choosing two points on the line). Let me calculate this ... What is the slope? *If we say M=Δy/Δx, the change in x, I will not take unit by unit, when I choose 10, the change in y, is like 8... so the slope is about 0.8."*

As the data suggested, he firstly chose two points on the placed informal line of best fit. Then, asking to himself as 'what is the slope', he calculated the *algebraic ratio* indicating the change in y variable with unit change in x variable.

Similarly, TE10 used algebraic ratio to find estimated slope value for 10th question:

"Let's say *it is 6 to 1,48. (6,1.48). Let's get another 7 here. That's better, if we do as such.. here is 1.42. 7 to 1.42 (7, 1.42) (calculating) m = 1.42-1.45 / 7-6 = -0.06 / 1."*

As the excerpts suggested, he chose two points on the placed informal line of best fit.Then, similar with TE9, he calculated the difference between x and y values of these points. In particular, since he expressed the slope of line of best fit as writing a ratio of changes in values of x and y variables, I categorized this interpretation as *algebraic ratio conception* of slope.

On the other hand, both teacher educators having signal conception used the slope value as *parametric coefficient* of the regression equation in conjunction with *functional property* conception.

Particularly, TE9 interpreted the slope value calculated for 9th question as follows:

"*0.7 is the coefficient. (substituting in y = mx + n) y = 0.7x + n.* This will be the slope. *Do you know what this is? The effect of variable x on y is 0.7***.** In other words, *it does not increase by 1 unit, it increases by 0.7. I mean from the drop height for each unit…If it is dropped from a height of 10 units, it will bounce by 7 units. If it is dropped from a height of 20 units, what does it do… there will be 14-unit effect."*

As the data indicated, he firstly thought to write a linear equation by using the slope value as a *parametric coefficient* of x variable. In particular, he substituted the slope value that he calculated as 0.7 on the linear regression equation. Then, he interpreted this slope value as rate of change between given contextual variables. In particular, he stated that increase in values of drop height for each unit would lead to increase in value of bounce height as 0.7 units. In this way, he referred to the *functional property* along with *real world situation conception*.

Similarly, TE10 also used functional property and algebraic ratio to interpret slope of the best fit line. In particular, for 10th question, he stated that:

"It's already negative. That would be such *as y = -0.06x + n.* Do you know what that means? *A 1-unit increase describes a negative decrease. That means as ticket prices increase, we decrease the number of people by 6%."*

As the excerpt showed, he firstly thought to substitude slope value as a *parametric coefficient* of x variable on the linear regression equation. Then, he mentioned the slope value as the rate of change between variables based on the *functional property*. In particular, he evaluated the change in y values with unit change of x values. By doing this, he described the slope value as the magnitude of the covariation of given variables. Also, he interpreted given functional situation using context variables. Therefore, he also referred to the *the real world situation* conception of slope. Lastly, since he indicated the behaviour of dataset as showing increase or decrease of the y variable with the unit change of x variable, he used the slope value regarding the *behaviour indicator conception*.

In particular, while these teacher educators having signal conception (TE9 and TE10) dominantly used the *algebraic ratio* to calculate the estimated slope values, they interpreted it using *functional property, parametric coefficient, behaviour indicator* and *real world situation*. In particular, similar to the teacher educators having predictor conception, they also interpreted the slope values as the rate of change between variables in given contextual information generally. In addition, when they were asked for the

comparison between slope meanings for linear functions and the line of best fit, TE10 stated that the slope value of the best fitting line is an *approximation* since the data points were not given on the same linear direction as were linear functions. This interpretation of slope might indicate similarity with his *signal conception* that appreciates variation in data set.

Considering all the results regarding slope conceptualization, teacher educators having *representer* conception (TE1, TE4, TE6, TE7, TE8) for informal line of best fit dominantly used the slope conceptions as *linear constant* and *parametric coefficient* along with *real world situation* and *behaviour indicator* conceptions. In particular, their interpretations regarding visual perspective of slope concept were predominant in compatible with their representer conception for the line of best fit. On the other hand, teacher educators having the *predictor* conception (TE2, TE3, TE5 and TE11) for the best fit line dominantly referred to the slope conceptions as *functional property* along with the *real world situation* and *behaviour indicator*. These participants generally elaborated on the slope meaning focusing on *predicted* or *approximate* rate of change between given contextual variables. This interpretation of slope is also parallel with their use of the line of best fit as finding predicted values for variables that are not included in dataset. Lastly, both teacher educators having the *signal* conception (TE9 and TE10) dominantly referred to the slope conceptions as *algebraic ratio, functional property*, *parametric coefficient, real world situation* and *behaviour indicator*. Particularly, these participants interpreted the slope meaning mostly focusing on the analytic perspective. In this respect, they generally used the functional structure of the line of best fit to calculate the slope value. However, they interpreted this value for the best fit line as an *approximation* since the data points are not lined up as if they are in linear functions. Hence, this approach also goes along with their signal in noise conception for the line of best fit.

6. DISCUSSION AND CONCLUSION

In this study, I investigated the subject matter knowledge of eleven mathematics teacher educators regarding the informal line of best fit. I specifically examined what conceptions and placement criteria for the informal line of best fit teacher educators had. Their slope conceptualizations regarding the informal line of best fit were also scrutunized. Considering the frequency of the mathematics teacher educators' reference to the relevant conceptions and criteria, dominant conceptions for each teacher educator were determined. Results showed that five of the teacher educators had *representer conception*, four of them had *predictor conception* and two of them had *signal conception* as their dominant conceptions, although none had neither *typical* nor *model* conceptions. Results also indicated that seven teacher educators used *closest* criterion, five of them referred to the *sum deviation* and two teacher educators also used *ellipse & area criterion* for the placement of the informal line of best fit dominantly. In addition, results regarding the slope conceptualizations showed that, almost all teacher educators had *behaviour indicator* and *real world situation conceptions* of slope dominantly such that they interpreted the slope as general increase-decrease in the data and were able to explain the slope given the context of the tasks. Besides these conceptions of slope, teacher educators holding representer *c*onception had also *parametric coefficient* and *linear constant conception* of slope, teacher educators with predictor conception also had *the functional property* conception of slope; and, teacher educators having the signal conception were also able to interpret the slope as the *algebraic ratio, functional property* and *parametric coefficient* dominantly.

In the following paragraphs, I elaborate on how the study addressed each of the research questions respectively. In the light of previous research, I discuss the results in detail and also point to the implications for further research and teaching.

6.1. Conceptions and Criteria

In this section, in order to determine the subject matter knowledge of teacher educators, I report on the results regarding the research questions "What conceptualization for informal line of best fit do teacher educators have?" and "What criteria do teacher educators use for placement of informal line of best fit?". First, I report on the general results; then, I discuss each conception in relation to the criteria used by teacher educators by pointing to the reasons as to how and why they utilized those criteria.

Previous research on conceptualizations of informal line of best fit has shown that both students and teachers had *model, typical, representer* and *predicto*r conceptions, while teachers also had *signal* conception (Casey, 2015a,b). Also, the results pointed that *representer conception* was the most common conception among students and teachers (Casey (2015a,b). Similar to Casey's studies, in this study, *representer conception* was also the most frequent conception used by teacher educators. Though, mathematics teacher educators depicted only the *signal, representer* and *predictor* conceptions dominantly. In particular, each teacher educator had more than one conception throughout the interview; but, TE1, TE4, TE6, TE7 and TE8 dominantly used the *representer*; TE2, TE3, TE5 and TE11 dominantly used the *predictor* and TE9 and TE10 dominantly used the *signal conception*. Also, TE2 used the *signal* conception along with the *predictor* conception.

Regarding the criteria for the placement of the line of best fit, previous research has shown that both students and teachers used *equal number* and *closest* criteria while the students additionally used *through as many points as possible* and *reflecting context* dominantly, and some teachers additionally applied *pairs* and *sum deviation* (Casey, 2015a,b). In this study, teacher educators also dominantly used the *closest* and *sum deviation* criterion as parallel with teachers. Different from previous research results, two of the teacher educators have dominant criteria as *ellipse & area approach* additionally. In particular, TE1, TE2, TE3, TE5, TE6, TE8, TE11 dominantly used the *closest* criterion, while TE2, TE7, TE8, TE9, TE10 applied *the sum deviation* and also TE2 and TE4 used the *ellipse and area approach*. In addition, different from previous research, data allowed

to determine the relationship between the participants' conceptions and the criteria they used. In the following paragraphs, I point to such relationship.

Previous research has shown that students having *representer conception* defined the informal line of best fit as a line that *'shows what the data looks like'* (Casey, 2015a) while teachers referred to it as *'the best representation of given dataset on the scatterplot'* (Casey, 2015b). In this study, similar to teachers, teacher educators also focused on the data on the scatter plot as a whole and emphasized that the most important aspect of the process is to represent the data. This indicated that they focused on the all data points at hand showing an aggregate view (Bakker, 2004; Biehler *et al.*, 2018) and fit the line to reach the best representation of the given dataset on the scatterplot. Results also pointed that such conceptualization afforded teacher educators with ways while placing the line of best fit such that they handled all the sample data within the scatter plot rather than focusing solely on particular data points. In particular, they mostly used the *the closest criterion* regarding the minimization of distances between all the data points and the best fit line. Also, taking into consideration all the data points, they balanced deviations above and below the line suggesting that they used *the sum deviation criterion*. The only exception was TE4 having the representer conception but using *the area approach* while placing the informal line of best fit for all the task situations. Area approach included drawing one (or more) convex shape surrounding all data points as closely as possible and almost equally dividing the shape to place the line of best fit as a representer of the data. Although different from the closest and the sum deviation criteria used by other participants with *representer conception*, such reasoning also reflects the common perspective among them as handling data from the aggregate view since all data points are taken into consideration. Additionally, I propose that it might be an efficient way of understanding the direction of association, strength of correlation and also making closer approximations for the informal line of best fit. That is, whenever the convex shape surrounding the whole data is drawn, it might provide opportunity to reason about the positive or negative association between variables. Similarly, the tickness of the convex shape might also allow for determining the strength of the association. That is, the tinner the convex shape is, the more likely it shows that the best fit line exists there.

Comparing these results with previous research done with students, an important difference seems to arise: Previous research has shown that students referred *'through as many points as possible'* or '*equal number of points on both sides'* criteria dominantly. Students' referring to the former criteria suggested the reflection of their previous knowledge on linear functions and their referring to the latter criteria reflected their use of average meaning in univariate setting limitedly. These altogether showed students' lack of insight in articulating the reasoning behind their criteria of placement (Sorto *et al.*, 2011) and also pointed to their case-oriented view (Casey, 2015a). On the other hand, teacher educators in this study dominantly used the *closest* and the *sum deviation* criteria. Additionally, they were able to explain the reasoning behind the use of their placement criteria injuxtapostion with their conceptions. Particularly, most teacher educators focused on the data at hand as a whole while making their placements. This is important because treating the data in aggregate view together with these two criteria allows for understading the reasoning behind the formal line of best fit (Sorto *et al.*, 2011).

Regarding the results *pointing to predictor conception*, previous research had identified only two teachers out of nineteen with this conception (Casey, 2015b). Similar to the explanations provided by teachers in earlier research, in this study, the participants using the *predictor conception* defined the line of best fit as '*a line that gives closer predictions for values not included in dataset'*. In addition, they mostly stated the purpose of using the linear regression as making predictions for future or unknown values. Their reasoning differed from the representer conception in some important ways: First, after determination of the best fitting line for the given dataset, they focused on this line and line equation to make predictions about the data that were not in the original data set (Casey, 2015a). Second, they had expectancy for strong association between variables since they thought that the stronger relationship would be resulted in the closer predictions. Moreover, they were able to examine the dispersion of data points in the original data set to interpret the variability in the predicted data. That is, the more deviated the values in the actual data, the more errors in predictions would be. All these suggested that teacher educators took into consideration of the original data set as a whole to further make inferences about the predicted values. Conference Board of Mathematical Sciences emphasized that not only teachers' but also students' knowledge of making "...informal treatment of inference..." from the given data set with respect to the predicted values is

important (CBMS, 2012, p.58). Such conceptualization also afforded teacher educators while placing the line of best fit: Particularly, regarding the placement criterion, compared to the teacher educators with representer conception, the participants having *predictor conception* mostly used the *closest criterion* for placement. Their reasoning behind the use of such criterion was to enable minimum deviations from the line of best fit to make predictions as close as possible to actual values. The only exception was TE2 having the predictor conception but using *the term ellipse* while placing the informal line of best fit for Task 5- 6- 7 including weakly associated variables. In particular, his approach was similar to the area approach used by TE4 such that he drew a shape around the data calling it as ellipse to identify the direction of association and the best fitting line for more dispersed data points as in Task 5-6-7. In this respect, he placed the informal line of best fit as passing through the endpoints of the ellipse and so divided it symmetrically. In other words, the data in these tasks were more dispersed from each other. Therefore, to control the deviations in the predicted values, he used the ellipse criterion while determining the best fit line as close as possible to all points.

Previous research has shown that two out of nineteen teachers regarded the line of best as the signal for data that eliminates inherent noise (e.g., the variability in the data) (Casey, 2015b). Results in this study pointing to *signal conception* depicted that three out of eleven teacher educators, including TE2, used this conception. Particularly, the teacher educators dealt with the variation contained in the given data set to extract a line that closely approximates the actual data values. This suggested that similar to previous research, teacher educators having this conception referred to the process of finding the best fitting line as '*eliminating the inherent noise'* (Casey, 2015b, p.17) to reach the signal in the data. That is, they realized that data included variation from linearity to some extent in normal and social circumstances such that this variation might stem from other affecting factors on the association between two variables. Therefore, the idea of data analysis is to eliminate inherent noise arose with other factors and obtain the signal -the best fitting line- as an approximation to the actual true relationship. This aligns with what Engel *et al.* (2008) stated such that variation is the main reason of applying statistical methods with the aim of filtrating signals from noisy data. In other words, the aim of the process is to extract the line of best fit showing and determining the order/stability in the data varied to some

extent. This stability refers to the certainty in circumstances including uncertainty (Konold & Pollatsek, 2002). As the data showed, teacher educators not only thought of the residuals while placing the line of best fit, but also interpret association between the variables constituting the line of best fit taking into consideration of the residuals. This aligns with previous research suggesting that the *signal* in the regression analysis of the bivariate data is the *best fit model* and *the noise* could be the *residuals* around it (Data= model fit+ residuals) (Borovcnik, 2005). This is paramount because researchers defining data analysis as a search for signal in noisy process (Konold & Pollatsek, 2002; Groth, 2005; and Engel *et al.*, 2008) emphasize the importance of the awereness of the actual data including some discrepancies from the determined model fit (Engel *et al.*, 2008). This is also important because the functional structure derived with the data analysis method (e.g., the formal line of best fit) should be interpreted with appreciation of variation based on the *signal meaning* (Engel *et al.*, 2008). Researchers further advocated that mathematics teachers and teacher educators who take the responsibility of statistic instruction should have the signal meaning (Konold & Pollatsek, 2002). In particular, organizing the instruction focusing on the *signal meaning* would be effective to improve students' statistical thinking and provide them to bridge between the deterministic view and the stochastic approach (Engel & Sedlmeier, 2011). Additionally, each participant having the *signal conception* for the meaning of the best fit line used the *sum deviation criterion* dominantly while placing the line. In particular, data suggested that since they aimed to control the inherent variation to reach the signal in the given dataset, the sum deviation criterion seemed to be supportive as focusing on the balance of deviations on each side of the best fitting line.

Considering all these conceptions injuxtaposition to each other, while *representer conception* of the informal line of best fit allows visually focusing on the best representation of the data points particular to the situation at hand, the *predictor conception* allows to utilize the line of best fit focusing on the functional structure beyond the given dataset to predict the future and unknown values of the variables. In addition, the predictor conception allows for examining the dispersion of data points in the original data set to interpret the variability in the predicted data. On the other hand, the *signal conception* allows simultaneously and holistically to focus on the data points particular to the situation, the predicted values and the residuals. Therefore, I contend that *signal conception* reflect

and include the ideas of both *representer* and *predictor conceptions*. Similarly, having the *representer* conception allows to focus on minimizing the distances between all the data points and balancing the deviations above and below the line while placing the best fit line. In addition, having the *predictor* conception allows taking into consideration of the minimum deviations from the line of best fit in order to make predictions as close as possible. Finally, having *signal* conception further allows focusing on the balance of deviations on each side of the best fitting line in order to control the inherent variation to reach the signal in the given dataset. Therefore, results suggested that all these conceptions go along with the two criteria, namely, *closest* and *sum deviation* with representer conception using both, predictor conception using the *closest* and signal conception using *the sum deviation* more dominantly.

In terms of statistical subject matter knowledge, all of the teacher educators' closely displacing the line of best fit might be considered common and comparable to what other professionals would know. Agreeing with Casey (2015a,b), such knowledgable act might be considered as the evidence of common content knowledge in statistical knowledge for teaching framework (Groth 2007, 2013). On the other hand, teaching mathematics require specialized mathematical knowledge not needed in other settings such that teachers need to know how to explain and justify the reasoning behind mathematical ideas (Ball, 2008). Therefore, teacher educators' mostly focusing on why and how they place the line of best fit the way they do and how they make sense of the placement criteria in terms of the reasoning behind the formal line of best fit with links to their conceptions might be taken as the evidence of their specialized content knowledge. Similarly, most of the teacher educators' being able to determine and explain the students' placement criteria in Task 8 referring both to their own criteria and to students' possible different criteria could be considered evidence of their specialized content knowledge (Groth, 2007, 2013). For instance, to judge students' line placements, the teacher educators examined the vertical distances between line and data points so that based on the closeness of the line to the data points they were able to determine that students used closest criterion. Similarly, based on the balance of residuals above and below the line, they were able to determine that students referred to the sum deviation criterion. Together with the earlier research focusing on students' limited placement criteria (e.g., equal number, through as many points as

possible) and teachers' limited placement criteria (e.g., equal number) (Casey, 2015a,b), I contend that teacher educators should utilize such knowledge of students/teachers combining with their knowledge of content to plan their instruction (Ball *et al.*, 2008).

6.2. Slope Conceptualizations for Informal Line of Best Fit

In this section, I report on the results regarding the research question "What slope conceptualizations do teacher educators use for conceptualization of informal line of best fit?" First, I report on the general results; then, I discuss each conception in relation to the slope conceptualization held by teacher educators by pointing to the reasons as to how and why they utilized such conceptualization of slope. In this respect, I elaborated on these conceptualizations in accordance with their conceptions of the best fit line to indicate implications for their relevance better.

Results showed that all teacher educators had dominantly *behaviour* and *real life conceptions* of slope such that they interpreted the slope as general increase or decrease in the data and were able to explain the slope based on the given context of the tasks. In addition, other frequently used slope conceptions were *functional property* (eight teacher educators), *parametric coefficient* (eight teacher educators), *algebraic ratio* (six teacher educators) and *linear constant* (five teacher educators) conceptualizations. Still, with respect to the three dominant conceptions (e.g. representer, predictor and signal), as further dominant slope conceptions, all teacher educators holding representer conception also had *parametric coefficient* and *linear constant conception* of slope, three teacher educators with predictor conception also had the *functional property* conception of slope; and, all teacher educators having the signal conception were also able to interpret the slope as the *functional property*, *algebraic ratio* and *parametric coefficient.*

Particularly, the participants (TE1, TE4, TE6, TE7, TE8) having *representer* conception for informal line of best fit dominantly referred to the slope conceptions as *real world situation*, *behaviour indicator*, *parametric coefficient* and *linear constant*

conceptualizations through the interview process. Their interpretations generally indicated to the increasing, decreasing or horizontal trend of the association between variables for given contextual information. This suggested their making sense of slope as *behaviour indicator* along with the *real world situation conceptions*. Also, they used the slope value as substituting on the regression equation for some tasks and questions. Therefore, they also referred slope as a *parametric coefficient* of linear regression equation. In addition, *linear constant conception* was emerged only for the participants having *representer* conception for the line of best fit. Particularly, they focused on collinearity of some data points before they placed the informal line of best fit on some task situations. In fact, their tendency to identify these collinear data points stemmed from their search for linearity in the given data set. However, since they were aware of the variability of data and described the regression line as the best representation of *whole* dataset, they used the closest criterion to find the best fit line instead of placing a line through a subset of data points. That is, for their placement of the line, they were not trapped into the reasoning some students held in previous research (Casey & Nagle, 2016). Particularly, earlier research showed that some students referred to the *linear constant conception* using a subset of data points during the line placement process. In contrast to how teacher educators reasoned, they tried to place the line passing through only those some data points and did not make any explanation regarding the use of closest criteria (Casey & Nagle, 2016). Again this might have been because of teacher educators' taking into consideration of the variability within the dataset as a whole. Therefore, it is important for teacher educators and teachers to take attention to the variability within the data set as well as thinking of the data as a whole (Konold & Pollatsek, 2002; Engel *et al.*, 2008). In addition, researchers suggested students to think of "...the linear constant conceptualization not as a description of the straightness of the points but as a description of the straightness of the model for these points" (p.175). I contend that such reasoning is important for students to identify that slope value of regression line is a constant property for the *predicted* values on the line of best fit. In fact, when the comparison between slope meanings for linear functions and the line of best fit was questioned during the interview, TE10 having the signal conception stated that the slope value of the best fitting line is an *approximation* since the data points were not collinear as were linear functions. Thus, this perspective might promote their understanding of slope in statistical setting as an *approximation* for the magnitude of the *actual* relationship between variables (Casey & Nagle, 2016). In addition, this knowledge

might provide making connection between linear functions and the linear regression subjects, and so deterministic and stochastic perspectives.

Similar to the participants having the representer conception, participants having the *predictor* conception (TE2, TE3, TE5 and TE11) for the best fit line referred dominantly to the slope meanings as *behaviour indicator, and real world situation;* but, differed from them by dominantly using the *functional property conception* and not making any reference to *linear constant* conception. In particular, the participants who referred to *functional property* (TE2, TE3 and TE5) along with *real world situation* interpreted slope values as an *approximate rate of change between context variables* for the given dataset varying to some extent. Similarly, Casey and Nagle (2016) also identified that students using *functional property* and *real world situation* conceptions of slope made more accurate line placements and interpretations. In addition, Casey (2015b) advocated having this *flexible perspective* for interpretation of slope value in the given context for teachers responsible for teaching statistics. Particularly, from the statistical point of view, since the dataset includes inherent variation, slope value should be interpreted as *"a typical, average or expected change and association between two variables"* (Casey, 2015b, p.26). In addition, since slope concept indicates the magnitude of association between variables, making context-specific interpretations based on this value would provide more meaningful understanding. On the other hand, teacher educators' use of the *functional property* further pointed to their making sense of the slope value from an analytic perspective rather than a visual perspective. In other words, they interpreted slope meaning focusing on the *functional property* that can be calculated by using the difference- quotient ($\frac{Y2-Y1}{X2-X1}$), or used as the coefficient of x in the equation form y=mx+n. This is in contrast to the visual perspective that refers to slope as a property of the graph of a function (Zavlavski et al., 2002).

Lastly, teacher educators (TE9, TE10) having the *signal* conception referred to the same categories of slope conceptions as *functional property, real world situation* and *behaviour indicator* as the participants having predictor conception. Particularly, they interpreted *functional property* as indicating covariation between variables, *behaviour indicator* as describing the trend of association and *real world situation* as showing the relationship of context variables. Different from the participants with the predictor

conception, these participants also used dominantly the *algebraic ratio* to calculate the estimated slope values by specifically using the difference quotient ($\frac{Y2-Y1}{X2-X1}$) for relevant questions and *parametric coefficient* as the coefficient of x in the equation form as y=mx+n to write linear regression equation.

Additionaly, using the slope of the line of best fit for making statistical inferences whether there is an association between variables could be considered as beyond describing the association between variables with the informal line of best fit. In this respect, testing whether slope value statistically significantly different from zero is used for more advanced inference tests for association (Kleinbaum, 1998). Therefore, Casey (2015b) suggested that interpreting bivariate data with no association by using zero-sloped line could be considered as mathematics teachers' horizon knowledge for teaching informal line of best fit. Considering mathematics teacher educators as the teachers of teachers, I contend that interpreting bivariate data with no association by using zero-sloped line might be within the terrain of specialized content knowledge for teacher educators. In particular, in this study, Task6 and Task7 were analyzed to understand teacher educators' perspectives in case of having dataset with nonassociated variables. In the previous study conducted with teachers, five teachers used horizontal line to indicate no association between variables on Task 6 (Casey, 2015b). Similarly, in this study, while five teacher educators (TE2, TE3, TE4, TE5, TE9) placed zero-sloped line for at least one task organized with nonassociated data and four of them (TE2, TE3, TE5, TE9) explained their horizontal line placement with reference to the nonassociation between variables. On the other hand, TE4 determined horizontal line for these tasks with the beneft of his placement criteria as area approach rather than consideration of nonassociation. Further, TE5 with predictor conception interpreted the horizontal line of best fit as indicating average of values for interest variable. In addition, TE5 also associated the slope and correlation coefficient values while elaborating on the case of nonassociation. For him, the case of nonassociation between variables would lead to have correlation coefficient (r) as equal to zero and so it requires horizontal line of best fit with zero-slope. On the other hand, modelling bivariate data with horizontal line of best fit (y=a) is also used in linear functions in mathematical setting (Casey, 2015b) Therefore, knowing the relationship between the case of nonassociation and horizontal line is essential for teachers (Casey,

2015b) and also teacher educators to reach necessary conclusions for association between variables and also provide connection between mathematical and statistical contents.

Results regarding conceptions held by teacher educators injuxtaposition to the results from their slope conceptualizations of the informal line of best fit point to the following: Results regarding the conceptions pointed that while *representer conception* allows visually focusing on the best representation of the data points particular to the situation at hand, the *predictor conception* allows to utilize the line of best fit focusing on the functional structure beyond the given dataset to predict the future and unknown values of the variables as well as examining the dispersion of data points in the original data set to interpret the variability in the predicted data. Results also pointed that the *signal conception* allows focusing on the data particular to the situation, the predicted values and the residuals simultaneously and holistically. Similar to these results, teacher educators having *predictor* conception and also *signal* conception interpreted the slope as referring to the functional property along with real world situation conception. Particularly, they interpreted slope as an *approximate rate of change between context variables* for the given dataset varying to some extent. This suggested that they were aware of the fact that slope referred to *"a typical, average or expected change between two variables"* (Casey, 2015b, p.26). Researchers also claimed that such view of slope is based on analytical perspective in contrast to the visual perspective that refers to slope as a property of the graph of a function (Zavlavski et al., 2002). Additionally, some teacher educators having *predictor* and *signal* conception made interpretation of nonassociation in relation with horizontal line of best fit. This interpretation indicates that they utilized the slope of line of best fit to make inference beyond to make description for association between variables. In contrast, the participant having representer conception who placed horizontal line of best fit explained his reasoning referring to his placement criterion as area approach.

In terms of statistical subject matter knowledge, I contend that all of the teacher educators' use of *behaviour indicator* and *real world situation* conceptions of slope while interpreting slope of the line of best fit might be considered common and comparable to what other professionals would know. Therefore, such knowledgable act might be considered as the evidence of common content knowledge in statistical knowledge for

teaching framework (Groth 2007, 2013). On the other hand, teacher educators having the *functional property* along with *real world situation* conceptions of slope might be considered as specialized content knowledge for teaching. This is because, having this *flexible perspective* for interpretation of slope value in the given context is important for teachers to have (Casey, 2015b). Also, from the statistical point of view, slope value should be interpreted as *"a typical, average or expected change and association between two variables"* (p.26) since the dataset includes inherent variation (Casey, 2015b).

7. IMPLICATIONS FOR RESEARCH AND TEACHING

Result of the study suggested that *signal conception* seem to reflect and include the ideas of both *representer* and *predictor* conceptions such that it is beyond those conceptions. Since, the results came from interview data, further research focusing on the teaching of the informal line of best fit needs to be conducted to support the hypothesis. Also, data came from eleven teacher educators; five with representer conception, four with predictor conception and two with signal conception. Therefore, the hypothesis might be further supported by evidence from a larger number of teacher educators representing each conception. Similarly, in this study, Tasks 1-7 used to investigate teacher eduators' interpretations regarding conceptions and criteria for placement of informal line of best fit. In particular, two of these tasks (Task1-2) reflected strong positive association, two of them showed strong negative association (Tasks3-4), one demonsrated moderate positive association with inclusion of an outlier point (Task5) and two of them represented nonassociation (Task 6-7) with different contexts. In this respect, the aforementioned hypothesis regarding the inclusion of representer and predictor conceptions within signal conception might be further investigated with use of larger number of tasks differentiating in contextual information, scatter of data points and strength of correlation. Further, results showed that teacher educators' informal line of best fit on nonassociated dataset (Task6-7) reflected more variability in placement compared to other task situation. This indicated that some teacher educators had difficulty to demonstrate nonassociated data with horizontal line of best fit and make interpretations relevant to nonassociation. Therefore, research examining the approaches on the tasks having more variability in the dataset might provide important implications to understand statistical inferences regarding association/ nonassociation. In particular, this knowledge also might promote teaching and learning of constant functions in mathematical setting.

Additionaly, data suggested a link between conceptions and the criteria used for the placement of the line of best fit. Teacher educators' conceptions of the informal line of best fit afforded their placement criteria such that they were able to explain how and why they made the placements. Though, earlier research pointed that students were not able to articulate the reasons behind their placement criteria. Further research focusing on how

students and teachers construct such connections need to be conducted for the development of their conceptualization of the informal line of best fit.

Results also suggested that teacher educators' conceptions were in relation with their conceptualization of the slope of the informal line of best fit. Further research needs to be conducted to further scrutinize the relationship among different conceptions of the informal line of best and the slope conceptualizations. Specifically, research conducted with students would have more importance since they have some conceptualizations of slope before being introduced to the line of best fit. Understanding possible associations between slope conceptions and conceptualization of line of best fit could promote more connected and coherent learning processes. On the other hand, understanding students' slope conceptualizations might also give important implications for teaching line of best fit in statistical setting as a continuation of linear functions in mathematical setting. Particularly, knowing students' slope conceptualizations and organizing lessons accordingly could enhance students' understanding of both linear function and the line of best fit. Lastly, bringing students to the idea of using the slope of line of best fit to make inference about association or nonassociation between variables might contribute to their learning of more advanced statistics. Therefore, research focusing on how students and teachers reach the understanding of making inference using the slope of the line of best fit need to be conducted for the development of their conceptualization of linear regression.

In addition, research in teaching and learning environment could give more insight into understanding teachers' and also teacher educators' conceptualizations of line of best fit. Therefore, conducting research on the teaching of the informal line of best fit might also shed light on the path towards understanding teachers' and teacher educators' statistical knowledge of content and students that might give more implications for teaching of this subject.

Lastly, in this study, results indicated that signal conception reflect and include the ideas of both representer and predictor conceptions of the line of best fit. In this respect, signal conception focusing on both functional structure and variability around it indicates stochastic approach to the linear regression. Therefore, organizing statistical education for

prospective teachers aiming to reach the knowledge of having signal conception might contribute their understanding and interpretation of association and regression. For this purpose, teacher educators' reasoning process and their articulations mentioned in this study might give important implications for planning the teaching and learning process. In addition, making use of similar tasks that utilized in this study to informally search for the line of best fit might promote prospective teachers' and students' understanding behind the procedural methods (e.g. least square regression method) and the purpose of implementing these methods.

REFERENCES

Association of Teacher Educators (ATE). (2006). *Standards for teacher educators*. Manassas Park, VA: Author.

Ball, D. L., Thames, M. H., and Phelps, G. (2008). Content Knowledge for Teaching: What Makes It Special? *Journal of Teacher Education*, 59(5), 389-407.

Bakker, A. (2004). Reasoning about shape as a pattern in variability. *Statistics Education Research Journal, 3*(2), 64-83.

Bakker, A., & Gravemeijer, K. P. (2004). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 147-168). Springer, Dordrecht.

Batenero, C., Estepa, A., Godino, J. D., & Green, D. R. (1996). Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education*, 27, 151-169.

Batanero, C., Godino, J., & Estepa, A. (1998). Building the meaning of statistical association through data analysis activities. In A. Olivier & K. Newstead (Eds.), Proceedings of the 22nd Conference of the International Group for the Psychology of Mathematics Education (pp. 221–236). Stellenbosch, South Africa: University of Stellenbosch.

Batanero, C., & Díaz, C. (2010). Training teachers to teach statistics: what can we learn from research? *Statistique et enseignement, 1*(1), 5-20.

Beswick, K., Goos, M., & Chapman, O. (2014, January). Mathematics teacher educators' knowledge. In *Cite as: Liljedahl, P., Nicol, C., Oesterle, S., & Allan, D.(Eds.).(2014). Proceedings of the 38th Conference of the International Group for the Psychology of Mathematics Education and the 36th Conference of the North American Chapter of the Psychology of Mathematics Education (Vol. 1). Vancouver, Canada: PME.* (p. 254).

Biehler, R., Frischemeier, D., Reading, C., & Shaughnessy, J. M. (2018). Reasoning About Data. In D. Ben-Zvi, K. Makar & J. Garfield (Eds.), *International Handbook of Research in Statistics Education* (pp. 139-192). Springer, Cham.

Borovcnik, M. (2005). Probabilistic and statistical thinking. In M. Bosch (Eds.), *Proceedings of the fourth congress of the European society for research in mathematics education* (pp. 485-507). Sant Feliu de Guixols, Spain.

Burrill, G., & Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. In C. Batanero, G. Burrill, and C. Reading (Eds.), *Teaching statistics in school mathematics-Challenges for teaching and teacher education* (pp. 57-69). Springer, Netherlands.

Carver, R., Everson, M., Gabrosek, J., Horton, N. J., Lock, R. H., Mocko, M., ... & Wood, B. (2016). Guidelines for assessment and instruction in statistics education: College Report 2016, (p.77). http://www. amstat. org/education/gaise, accessed in May 2019.

Casey, S. A. (2008). *Subject matter knowledge for teaching statistical association*. Doctoral Dissertation, Illinois State University.

Casey, S. A. (2010). Subject matter knowledge for teaching statistical association. *Statistics Education Research Journal*, *9*(2).

Casey, S. A. (2014). Teachers' knowledge of students' conceptions and their development when learning linear regression. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in Statistics Education: Proceedings of the Ninth International Conference on Teaching Statistics, Flagstaff, Arizona, USA*. Voorburg, The Netherlands: International Statistical Institute.

Casey, S. A. (2015a). Examining student conceptions of covariation: A focus on the line of best fit. *Journal of Statistics Education*, *23*(1).

Casey, S. A., & Wasserman, N. H. (2015b). Teachers' Knowledge About Informal Line of Best Fit. *Statistics Education Research Journal*, *14*(1).

Casey, S. A., & Nagle, C. (2016). Students' use of slope conceptualizations when reasoning about the line of best fit. *Educational Studies in Mathematics*, *92*(2), 163-177.

CBMS (2012). *Conference Board of Mathematical Sciences*.

Clement, J. (2000). Analysis of clinical interview: foundations and model viability. A. E.Kelly, R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education.* (pp. 547-589). London: Lawrence Erlbaum Associates Publishers.

Cobb, P., McClain, K., & Gravemeijer, K. (2003). Learning about statistical covariation. *Cognition and instruction*, *21*(1), 1-78.

Coladarci, T., Cobb, C. D., Minium, E. W., & Clarke, R. C. (2010). *Fundamentals of statistical reasoning in education*. Hoboken, John Wiley & Sons.

Common Core State Standards Initiative. (2010). Common core state standards for mathematics. http://www.corestandards.org/assets/CCSSI _Math% 20Standards.pdf, accessed in May 2019.

Çetinkaya, B., Erbaş, A. K., Çelikdemir, K., Koyuncu, F., Kol, M., Alacacı, C. (2017). Professional Competencies That Mathematics Teacher Educators Should Have: Reflections From a Workshop. IN: Galindo, E., & Newton, J., (Eds.). (2017). *Proceedings of the 39th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Indianapolis: Hoosier Association of Mathematics Teacher Educators.

Engel, J., Sedlmeier, P., & Wörn, C. (2008). Modeling scatterplot data and the signal-noise metaphor: Towards statistical literacy for pre-service teachers. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.), *Proceedings of the ICMI study 18 and IASE round table conference.* International Commission on Mathematics Instruction and International Association for Statistical Education: Monterrey, Mexico.

Engel, J., & Sedlmeier, P. (2011). Correlation and regression in the training of teachers. In *Teaching statistics in school mathematics-challenges for teaching and teacher education* (pp. 247-258). Springer, Dordrecht.

Estepa, A., & Sánchez Cobo, F. T. (2001). Empirical research on the understanding of association and implications for the training of researchers. In C. Batanero (Eds.), *Training researchers in the use of statistics*, 37-51. Granada: International Association for Statistical Education and International Statistical Institute.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2005). Guidelines for assessment and instruction in statistics education (GAISE) report. *Alexandria: American Statistical Association*.

Garfield, J., & Ben-Zvi, D. (2004). Research on statistical literacy, reasoning, and thinking: Issues, challenges, and implications. In D. Ben Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 397-409). Springer, Dordrecht.

Goldin, G. A. (2000). A scientific perspective on structured, task-based interviews in mathematics education research. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of Research Design in Mathematics and Science Education.* Lawrence Erlbaum Associates, Mahwah.

Gravetter, F. J., & Wallnau, L. B. (2013). Introduction to Regression. In *Statistics for the Behavioral Sciences* (pp. 558-572). Cengage Learning, Wadsworth.

Groth, R. E. (2005). An investigation of statistical thinking in two different contexts: Detecting a signal in a noisy process and determining a typical value. *The Journal of Mathematical Behavior*, *24*(2), 109-124.

Groth, R. E. (2007). Toward a conceptualization of statistical knowledge for teaching. *Journal for research in Mathematics Education*, *38* (5), 427-437.

Groth, R. E. (2013). Characterizing key developmental understandings and pedagogically powerful ideas within a statistical knowledge for teaching framework. *Mathematical Thinking and Learning*, *15*(2), 121-145.

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for research in mathematics education*, *39* (4), 372-400.

Henriques, A. C., & Oliveira, H. (2013). Prospective teachers' statistical knowledge for teaching when analyzing classroom episodes. In A.M. Lindmeier & A. Heinze (Eds.) *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 41-48). Kiel, PME.

Jaworski, B. (2008). *Development of the mathematics teacher educator and its relation to teaching development.* IN: Jaworski, B. and Wood, T. (eds.). The International Handbook of Mathematics Teacher Education Volume 4: The Mathematics Teacher Educator as a Developing Professional. Rotterdam: SensePublishers, pp. 335-361.

Kleinbaum, D. G., Kupper, L. L., Muller, K. E. & Nizam, A. (1998). *Applied Regression Analysis and Other Multivariable Methods*. Pacific Grove: Duxbury Press.

Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for research in mathematics education*, *33*(4) 259-289.

Konold, C., & Pollatsek, A. (2004). Conceptualizing an average as a stable feature of a noisy process. In D. Ben Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 169-199). Springer, Dordrecht.

Koponen, M., Asikainen, M. A., Viholainen, A., & Hirvonen, P. E. (2016). Teachers and their Educators-Views on Contents and their Development Needs in Mathematics Teacher Education. *The Mathematics Enthusiast*, *13*(1), 149-170.

Masingila, J. O., Olanoff, D., & Kimani, P. M. (2018). Mathematical knowledge for teaching teachers: knowledge used and developed by mathematics teacher educators in learning to teach via problem solving. *Journal of Mathematics Teacher Education*, *21*(5), 429-450.

Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative research: A guide to design and implementation*. John Wiley & Sons.

Moore-Russo, D., Conner, A., & Rugg, K. I. (2011). Can slope be negative in 3-space? Studying concept image of slope through collective definition construction. *Educational studies in Mathematics*, *76*(1), 3-21.

Moritz, J. (2004). Reasoning about covariation. In D. Ben Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 227-255). Springer, Dordrecht.

Nagle, C., Moore-Russo, D., Viglietti, J., & Martin, K. (2013). Calculus Students'and Instructors' Conceptualizations of Slope: A Comparison Across Academic Levels. *International Journal of Science and Mathematics Education*, *11*(6), 1491-1515.

Nagle, C., Casey, S., & Moore-Russo, D. (2017). Slope and Line of Best Fit: A Transfer of Knowledge Case Study. *School Science and Mathematics*, *117*(1-2), 13-26.

NCTM, N. (2000). Principles and standards for school mathematics.

Sorto, M. A., White, A., & Lesser, L. M. (2011). Understanding Student Attempts to Find a Line of Fit. *Teaching Statistics*, *33*(2), 49-52.

Stanton, M., & Moore-Russo, D. (2012). Conceptualizations of slope: A review of state standards. *School Science and Mathematics*, *112*(5), 270-277.

Statistics Education Research Journal, November, 2008.

Strauss, A., & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Sage Publications, Inc.

Stump, S. (1999). Secondary mathematics teachers' knowledge of slope. *Mathematics Education Research Journal*, *11*(2), 124-144.

Stump, S. L. (2001). High school precalculus students' understanding of slope as measure. *School Science and Mathematics*, *101*(2), 81-89.

Superfine, A. C., & Wagreich, P. (2009). Developing mathematics knowledge for teaching in a content course: A "design experiment" involving mathematics educators and mathematicians. *AMTE monograph series*, *6*, 25-42.

Team, C. C. W. (2011). Progression for the Common Core State Standards in Mathematics (draft). *Comment at commoncoretools. wordpress. com*, *1*, 111.

Tzur, R. (2001). Becoming a mathematics teacher-educator: Conceptualizing the terrain through self-reflective analysis. *Journal of Mathematics Teacher Education*, *4*(4), 259-283.

Zaslavsky, O., Sela, H., & Leron, U. (2002). Being sloppy about slope: The effect of changing the scale. *Educational Studies in Mathematics*, *49*(1), 119-140.

Zaslavsky, O. (2009). Mathematics educators' knowledge and development. In *The professional education and development of teachers of mathematics* (pp. 105-111). Springer, Boston, MA.

Zieffler, A. S., & Garfield, J. B. (2009). Modeling The Growth of Students' Covariational Reasoning During an Introductory Statistics Course. *Statistics Education Research Journal*, *8*(1).

Zieffler, A., Garfield, J., Delmas, R., & Reading, C. (2008). A Framework to Support Research on Informal Inferential Reasoning. *Statistics Education Research Journal*, *7*(2).

Zopf, D. (2010). *Mathematical knowledge for teaching teachers: The mathematical work of and knowledge entailed by teacher education.* Unpublished doctoral dissertation, University of Michigan, Ann Arbor, Michigan.

APPENDIX A: INTERVIEW PROTOCOL

Introduction: In just a few minutes I am going to give you six tasks to complete. Each one involves a scatter plot graph of data. We will talk about the data on each graph and then I will ask you to determine the line of best fit for the data points. I would like you to think out loud as you determine where the line is on each graph so that I can understand how you are deciding where to place it.

Task 1:

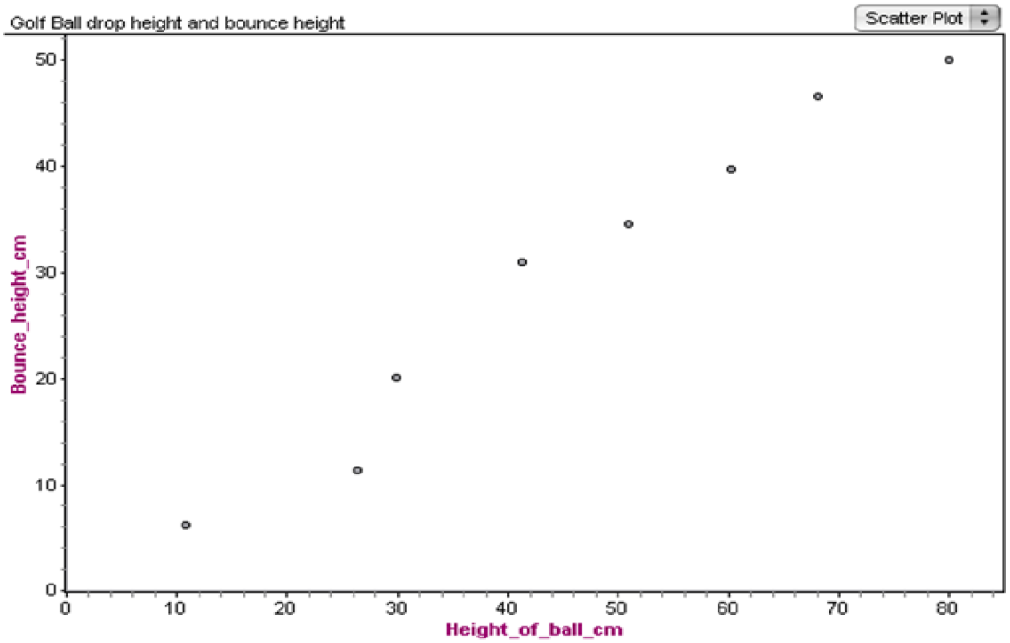

Figure A. 1. Golf ball bounce height and drop height.

Here is the first scatter plot which shows data about how high a ball bounces when it is dropped from different heights. For example, this point shows the golf ball was dropped from 10.8 centimeters and bounced back up to 6.1 centimeters. I would like you to use this piece of wire as the line and place it where it best fits the data. There are no right or wrong answers. I am interested in how you think about placing the piece of wire, so please tell me what you are thinking as you do this so I can follow your thoughts. Once you decide where

you want the line on the scatter plot, we are going to use scotch tape to keep the wire in place [Allow teacher to work through the task].

After each line of best fit is completed (if teacher did not mention when talking aloud), ask follow-up questions where applicable, including:

- Why did you choose to draw the line there?
- What criteria did you use in deciding where to draw the line?
- How did you decide on that criteria?
- Could you say more about how you decided at the end that it should be angled a bit more? *(If applicable)*
- *(If applicable):* I noticed you moved the line so that it went through here *(point to the intersection of the axes).* Can you talk about why you thought that was important?
- Can you talk about what the line shows about the relationship between bounce height of a ball and the height it was dropped from?

Task 2:

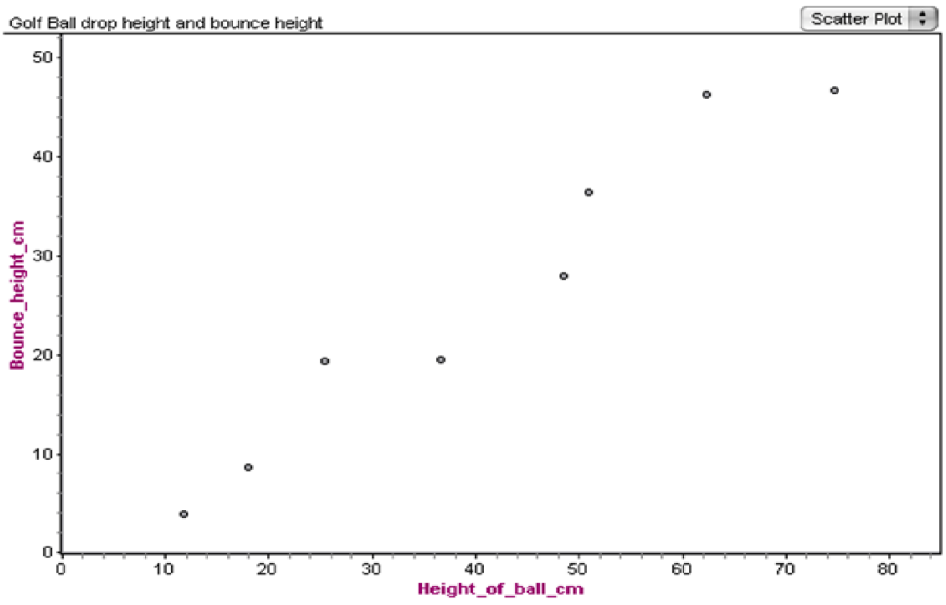

Figure A. 2. Golf ball bounce height and drop height.

Here is the second scatter plot which shows data about the same context in Task 1. I would like you to use the piece of wire as the line and place it where it best fits the data.

- Why did you choose to draw the line there?
- What criteria did you use in deciding where to draw the line?
- Did your criteria change for this graph?
- How did you decide on that criteria?
- Could you say more about how you decided at the end that it should be angled a bit more? *(If applicable)*
- *(If applicable):* I noticed you moved the line so that it went through here *(point to the intersection of the axes)*. Can you talk about why you thought that was important?
- The line you placed for task #2 is different in some ways than the line that you placed in task #1. Please talk about some of the ways they are different.

Task 3:

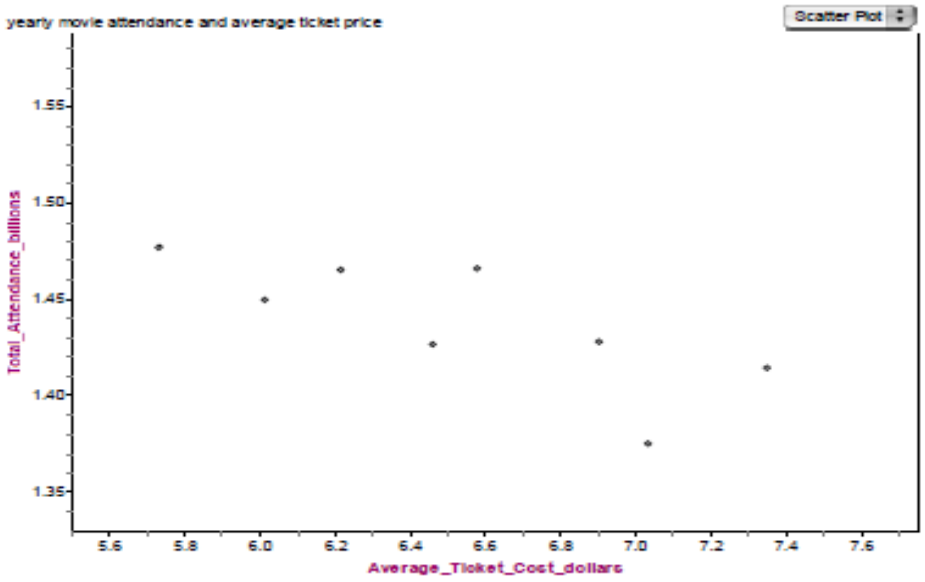

Figure A. 3. Average ticket cost and total attendance.

The data on this graph [Task 3] estimate the total number of movie tickets sold over an entire year and also estimate the average price of a ticket for that year based on sales in a sample of cities. For example, second point shows that during one year when the average price of a ticket was about $6 the total number of people attending movies was about 1.45 billion people. Where do you think the line of best fit would be for this data?

- Why did you choose to draw the line there?
- What criteria did you use in deciding where to draw the line?
- Did your criteria change for this graph?
- How did you decide on that criteria?
- Could you say more about how you decided at the end that it should be angled a bit more? *(If applicable)*
- Can you talk about what the line shows about the relationship between attendance totals for movies and the price of a movie ticket?

Task 4:

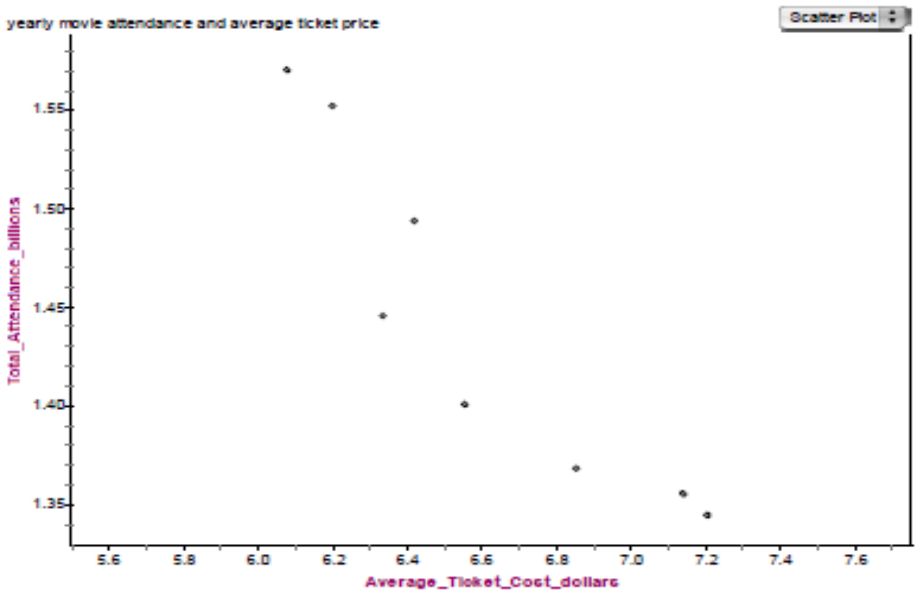

Figure A. 4. Average ticket cost and total attendance.

The next graph [Task 4] shows similar data [to Task 3] that is calculated using movie sales data from a different set of cities during the same year. Where do you think the line of best fit would be for this data?

- Why did you choose to draw the line there?
- What criteria did you use in deciding where to draw the line?
- Did your criteria change for this graph?
- How did you decide on that criteria?
- Could you say more about how you decided at the end that it should be angled a bit more? *(If applicable)*

Task 5:

The next scatter plot displays data from a different setting.

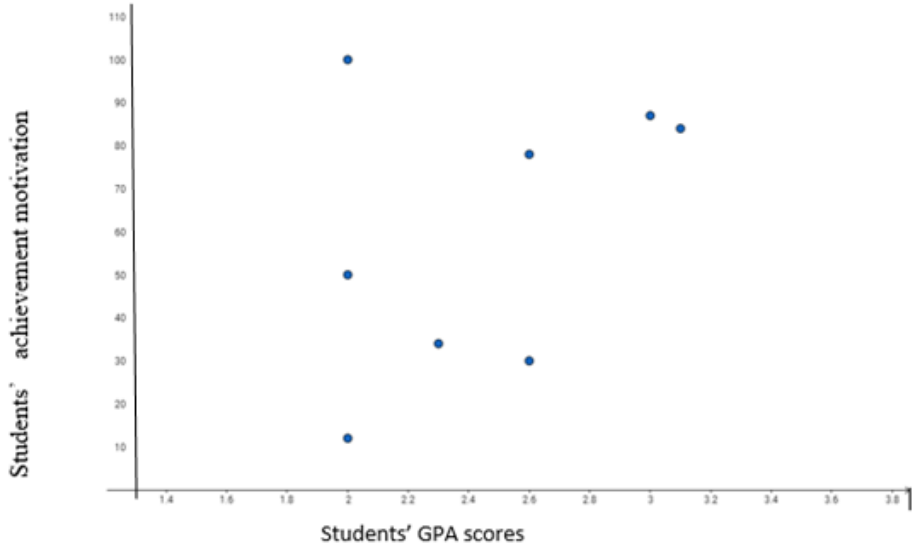

Figure A. 5. Students' achievement motivation and GPA scores.

This graph show data about students' achievement motivation and their GPA scores. GPAs can range from 0 to 4 and motivation scores in this example range from 0 to 100. For example, a student has a GPA score of 2 and his/ her motivation score is recorded as 12. Where do you think the line of best fit would be for this data?

- Why did you choose to draw the line there?
- What criteria did you use in deciding where to draw the line?
- Did your criteria change for this graph?
- How did you decide on that criteria?
- Could you say more about how you decided at the end that it should be angled a bit more? *(if applicable)*
- Can you talk about what the line shows about the relationship between students' achievement motivation and their GPA scores?

Task 6:

The next scatter plot displays data from a different setting.

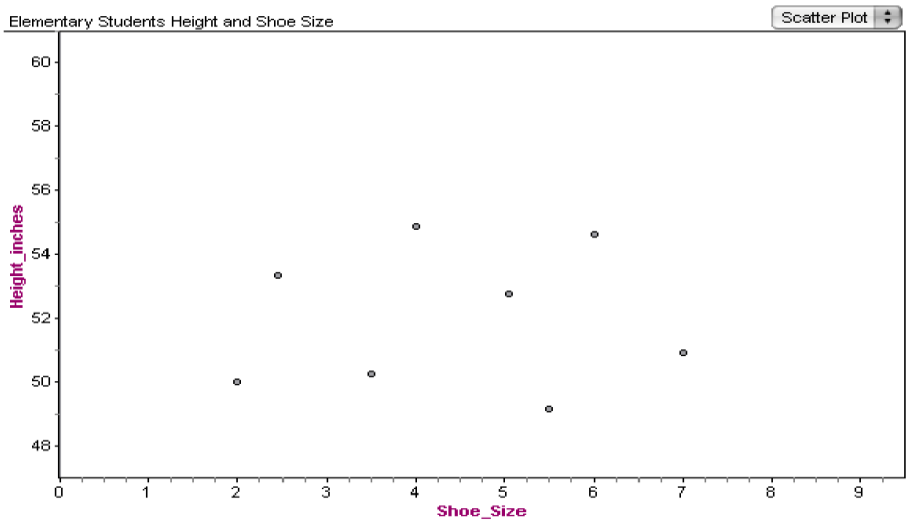

Figure A. 6. Students' height and shoe size.

An elementary teacher wrote down the height in inches and shoe size for each of 8 students in his class. He plotted the data in this scatter plot [Task 5]. Where do you think the line of best fit would be for this data?

- Why did you choose to draw the line there?

- What criteria did you use in deciding where to draw the line?
- Did your criteria change for this graph?
- How did you decide on that criteria?
- Could you say more about how you decided at the end that it should be angled a bit more? *(if applicable)*
- Can you talk about what the line shows about the relationship between elementary students' height and shoe size?

Task 7

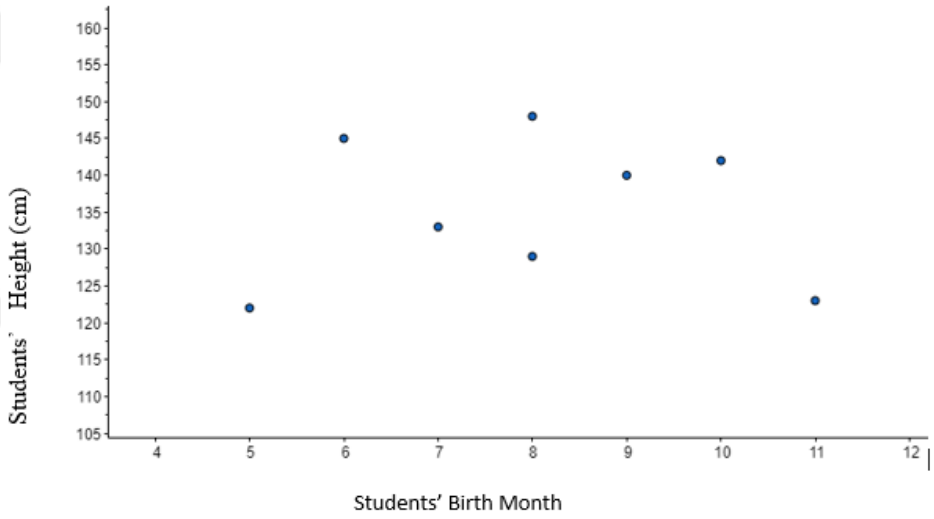

Figure A. 7. Students' height and birth month.

This graph show data about randomly chosen eight students' height (cm) and their birth month in an elementary school. Where do you think the line of best fit would be for this data?

- Why did you choose to draw the line there?
- What criteria did you use in deciding where to draw the line?
- Did your criteria change for this graph?
- How did you decide on that criteria?

- Could you say more about how you decided at the end that it should be angled a bit more? *(if applicable)*
- Can you talk about what the line shows about the relationship between elementary students' height (cm) and their birth month?

After Tasks 1-7, ask the following questions regarding the line of best fit:

- Could you tell me what you would say to a student that asked you "What is the line of best fit?"
- Could you tell me what you would say to a student to help them draw the line of best fit on a scatter plot?

Task 8:

Some students were given the same task you were given in task 1 (to find the line of best fit for the data on a golf ball's drop height and bounce height), task 3 (to find the line of best fit for the data on yearly movie attendance and average ticket price) and task 6 (to find the line of best fit for the data on elementary students' height and shoe size). They had different solutions for each task some of them are given below. In each case: which student's line fits the data better and why?

For Task 1

Angelo's solution:

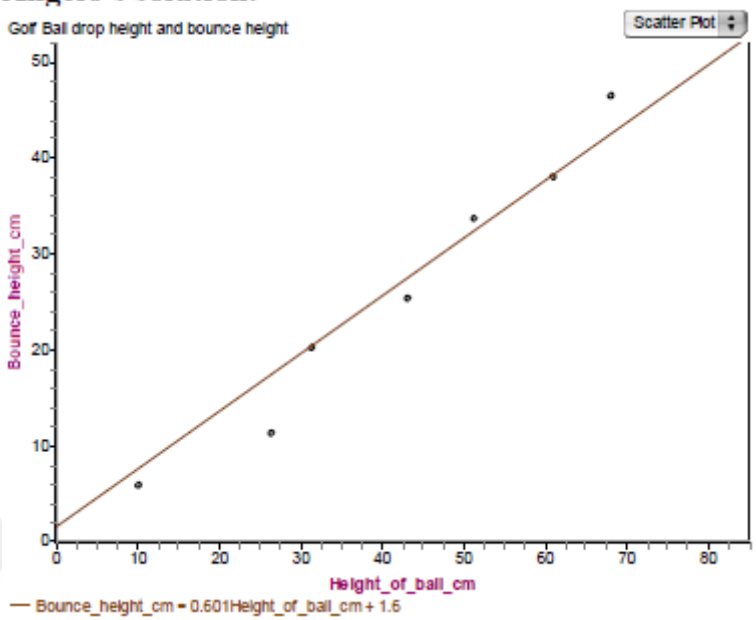

Figure A. 8. Angelo's line placement for Task1.

Barbara's solution:

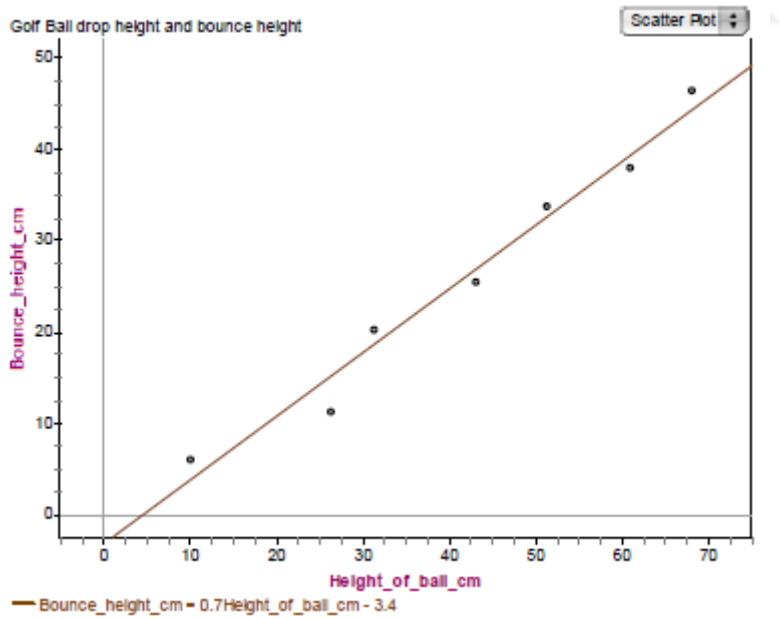

Figure A. 9. Barbara's line placement for Task1.

For Task 3

Angelo's Solution

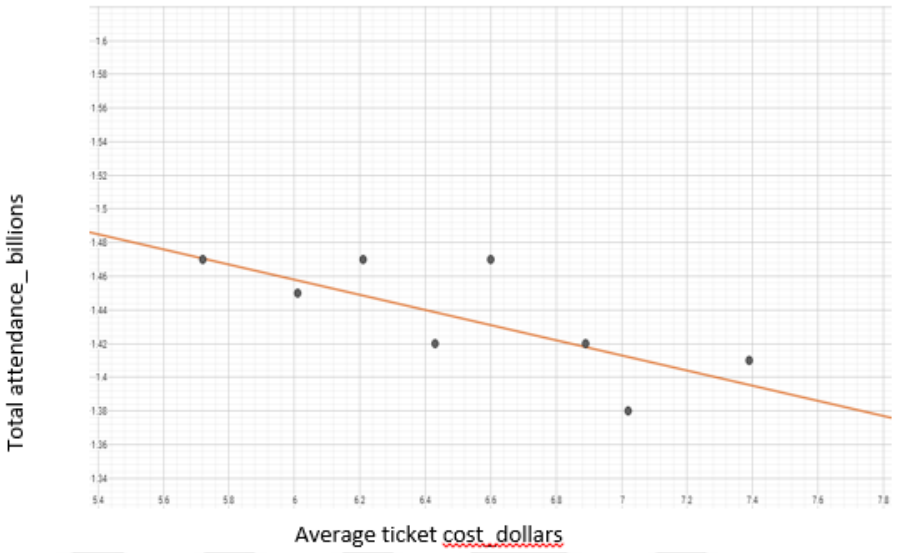

Figure A. 10. Yearly movie attendance and average ticket price.

Total attendance_billions= 1.728- 0.045average ticket cost_dollars

Barbara's Solution

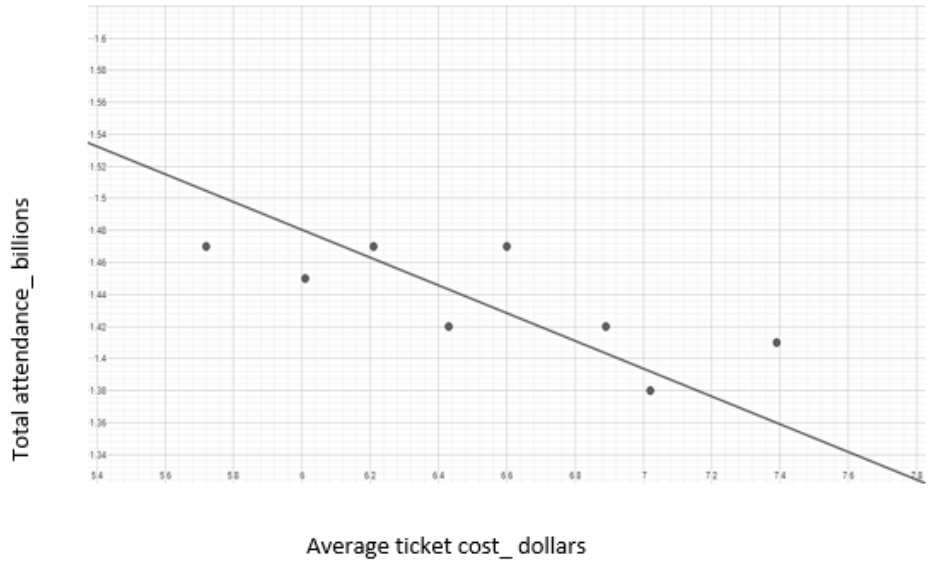

Figure A. 11. Yearly movie attendance and average ticket price.

Total attendance_billions= 2 - 0.09average ticket cost_dollars

For Task 6

Angelo's Solution

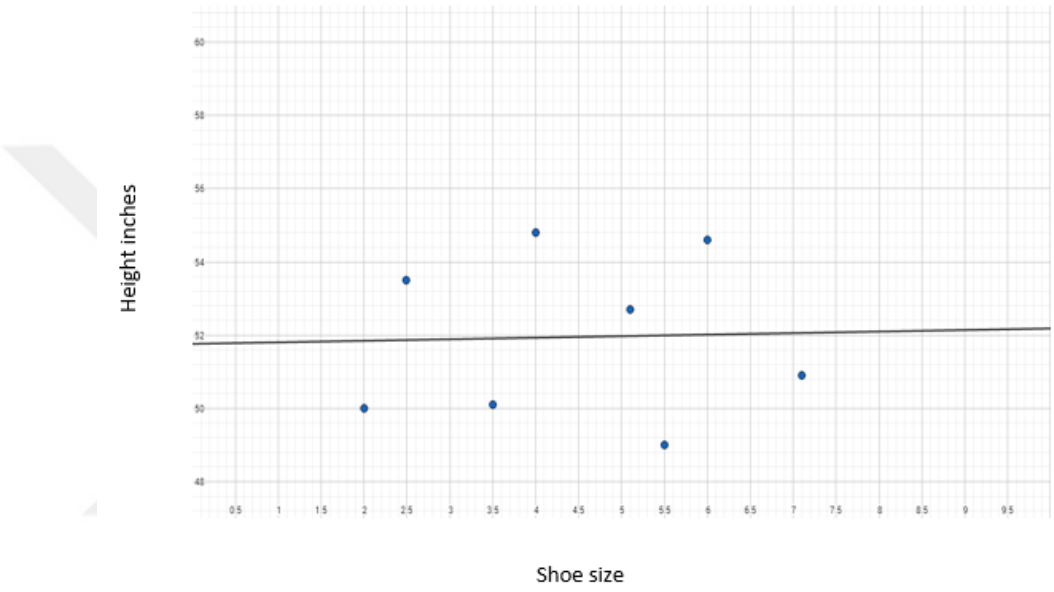

Figure A. 12. Elementary students' height and shoe size.

Height_inches= 51.761+ 0.042shoe size

Barbara's Solution

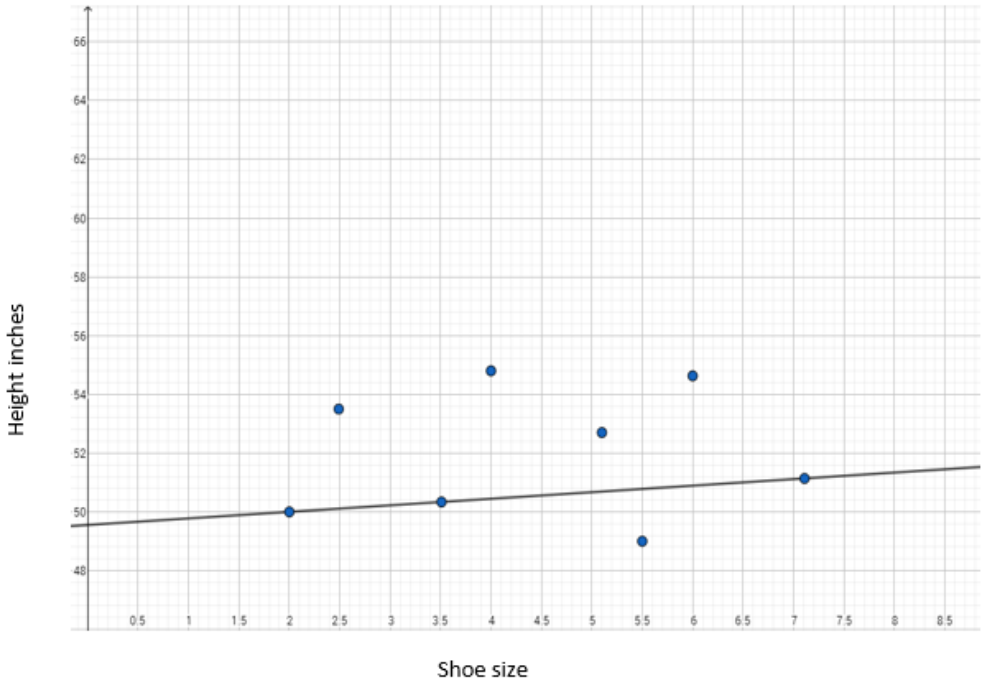

Figure A. 13. Elementary students' height and shoe size.

Height_inches= 49.557+ 0.222shoe size

Jane's Solution

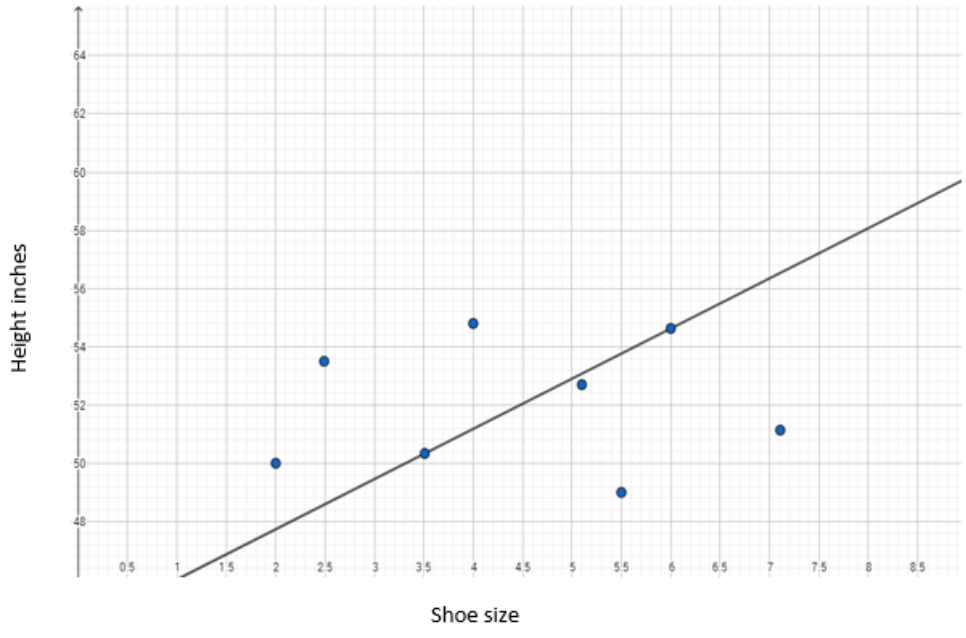

Figure A. 14. Elementary students' height and shoe size.

Height_inches= 44.29+ 1.723shoe size

9) *(Show teacher task #1 again)*

- Could you repeat your thoughts about what the line shows about the relationship between golf ball drop height and bounce height?
- Would you take a moment and calculate an estimate of the slope of the best fit line you determined? You may write at the bottom of the paper.
- Can you talk about the meaning of *(say the slope value here)* in the context of golf ball drop height and bounce height?

10) *(Show teacher task #3 again)*

- Could you repeat your thoughts about what the line shows about the relationship between attendance totals for movies and the price of a movie ticket?
- Would you take a moment and calculate an estimate of the slope of the best fit line you determined? You may write at the bottom of the paper.
- Can you talk about the meaning of *(say the slope value here)* in the context of movie attendance totals and ticket prices?
- Is there any difference between slope of linear functions and slope of line of best fit? If yes, how?