
GAZİ ÜNİVERSİTESİ
Gazi
1926

COMPREHENSIVE WRITTEN CORRECTIVE FEEDBACK: EFFECTS ON L2 SYNTACTIC ACCURACY AND COMPLEXITY IN WRITING

Mohammadreza Valizadeh

A PHD DISSERTATION
ENGLISH LANGUAGE TEACHING

GAZI UNIVERSITY
INSTITUTE OF EDUCATIONAL SCIENCES

JULY, 2019

COPYRIGHT AND CONSENT TO COPY THE DISSERTATION

All rights of this dissertation are reserved. It can be copied …… months after the date of delivery on the condition that reference is made to the author of the dissertation.

AUTHOR:

Name: Mohammadreza

Last name: Valizadeh

Signature

Date of delivery: May 29, 2019

DSSERTATION:

Title of dissertation in Turkish: "İkinci Dilde Yazmada Kapsamlı Yazılı Düzeltici Geri Bildirimin Sözdizimsel Doğruluk ve Karmaşıklık Üzerine Etkileri"

Title of dissertation in English: "Comprehensive Written Corrective Feedback: Effects on L2 Syntactic Accuracy and Complexity in Writing"

DECLARATION OF CONFORMITY TO ETHICS

I declare that I have complied with the scientific ethical principles within the process of typing the dissertation that all the citations are made in accordance with the principles of citing and that all the other sections of the study belong to me.

First and last name of the author: Mohammadreza Valizadeh

Signature of the author:

JÜRİ ONAY SAYFASI

We certify that the dissertation entitled “Comprehensive Written Corrective Feedback: Effects on L2 Syntactic Accuracy and Complexity in Writing” prepared by Mohammadreza Valizadeh has been unanimously found satisfactory by the jury for the award degree of doctorate of philosophy in the subject matter of English language teaching at Gazi University, department of foreign languages and education.

Supervisor Assoc. Prof. Dr. Hacer Hande Uysal Gürdal ……

Department of Foreign Languages Education, Hacetteppe University

Member Assoc. Prof. Dr. Nurdan Özbek Gürbüz ……

Department of Foreign Language Education, Middle East Technical University

Member Assoc. Prof. Dr. Kadriye Dilek Akpinar ……

Department of Foreign Languages, Gazi University

Member Associate Prof. Dr. Cemal Çakır ……

Department of Foreign Languages, Gazi University

Member Assistant Prof. Dr. Hatice Ergül ……

Department of Foreign Languages Education, Hacetteppe University

Date of dissertation defense: July 1, 2019

I certify that this dissertation has complied with the requirements of degree of Doctorate of Philosophy in the subject matter of English Language Teaching.

Prof. Dr. Selma Yel …….

Director of Institute of Educational Sciences

ACKNOWLEDGEMENTS

I would like to thank all the people who have helped me and offered me excellent advice while I completed my studies for the PhD education including this dissertation.

First, my heartfelt thanks go to Prof. Hacer Hande Uysal Gürdal for her kind support as well as insightful comments and guidance.

I am also profoundly grateful to Prof. Abdulvahit Çakır, the head of the foreign languages department at Gazi University. I would also like to express my sincere thanks to Prof. Paşa Tevfik Cephe, the head of the ELT program, who provided me with fruitful classes during my study at Gazi University.

I also greatly appreciate Dr. Murat Kaplan's help, the head of Preparatory School at Antalya International University, who permitted me to collect data at Antalya International University.

Last but not least, I would like to thank all the participants at Antalya International University in Antalya who helped me with the data collection phase of this study.

İkinci Dilde Yazmada Kapsamlı Yazılı Düzeltici Geri Bildirimin Sözdizimsel Doğruluk ve Karmaşıklık Üzerine Etkileri

Doktora Tezi

Mohammadreza Valizadeh

GAZİ ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Temmuz, 2019

ÖZ

Bir ön test-müdahale-son test-gecikmeli son test tasarımını kullanan yarı-deneysel olan mevcut çalışma iki farklı tipte, kapsamlı (odaklanmamış) yazılı düzelteci geri bildirimi stratejisinin etkilerini araştırmayı amaçlamıştır: Hem kısa vadede hem de uzun vadede doğrudan düzeltici geribildirim ve metadilbilimsel açıklamaları. Hazırlık okulunda yazma derslerine kayıtlı yaşları 18-23 arasında değişen ve İngilizceyi yabancı dil olarak öğrenen 90 Türk öğrenci çalışmanın katılımcılarını oluşturmuştur. Katılımcıların ikinci yabancı dil yeterlilik seviyelerini kontrol etmek ve bu seviyenin homojen olduğundan emin olmak için, birinci haftada, ilk olarak Hızlı Yerleştirme Testinin (HYT) kalem ve kâğıt versiyonu verilmiştir. Öğrencilerin dil seviyesine karar vermek için Geranpayeh'in (2003) rehberi kullanılmıştır. 60 üzerinden 40 - 47 arasında değişen puanlara sahip olan öğrenciler (yani, Geranpayeh (2003)' e göre üst-orta seviye), yazılı metinlerde sözdizimsel doğruluğu ve karmaşıklığı ölçmek için kullanılan IELTS Yazma Görev 2 testi için seçilmiştir. Bu bilgi, öğrencilerin yazılı doğruluk ve karmaşıklıkları bakımından homojenliğini sağlamak için kullanılmıştır. Bu bağlamda, homojen olan öğrenciler: yalnızca içerik, yazım ve organizasyon hakkında geribildirim veren dilbilgisi hataları konusunda görüş bildirmeyen kontrol grubu (n = 30), dilbilgisi hataları hakkında geri bildirim alan doğrudan düzeltici geribildirim (n = 30) ve metadilbilimsel açıklamaları (n= 30) olmak üzere üç gruba ayrılmışlardır. Kontrol süresi beş hafta sürmüştür. Bütün katılımcılar her hafta sınıfta tartışmacı bir kompozisyon yazmış ve onu teslim etmişlerdir; araştırmacı öğretmen her gruba belirtilen geri bildirimleri sağlamıştır. Daha sonra öğrencilerden düzeltilmiş metinlerini gözden geçirmeleri istenmiştir. 6. haftanın 1. oturumunda son test yapılmış; 6. ve 7. haftalarda yazı üzerine çalışma yapılmamıştır. 8. haftada, gecikmeli son test uygulanmıştır. Sözdizimsel doğruluğu değerlendirmek için, aşağıdaki formül kullanılmıştır: [sözdizimsel hataların toplam sayısı / toplam sözcük sayısı] x 100. Ayrıca, sözdizimsel karmaşıklığı hesaplamak için, "Web tabanlı L2SCA: Tek Mod" başlıklı Lu'nun (2010) web tabanlı L2 Sözdizimsel Karmaşıklık Çözümleyicisi kullanılmıştır. Bu çalışmada 14 endeksten 10'u analiz edilmiştir. Bu on sözdizimsel karmaşıklık ölçümleri şunları içeriyordu: ortalama cümlecik uzunluğu, ortalama cümle uzunluğu, T biriminin

ortalama uzunluğu, cümle başına cümlecik, T-birimi başına cümlecik, T birimi başına karmaşık T birimleri, cümlecik başına bağımlı cümlecikler, T birimi başına koordinat cümleleri, T birimi başına karmaşık nominal değerler ve T birimi başına fiil cümleleri. Sonunda, her iki deney grubu da kısa ve uzun vadede yazılı metinlerin sözdizimsel doğruluğu açısından kontrol grubundan anlamlı ölçüde daha iyi bir performans göstermiştir; doğrudan düzeltici geribildirim ve metadilbilimsel açıklamada hem kısa hem de uzun vadede anlamlı bir fark bulunmamıştır. Kapsamlı yazılı düzeltici geribildirim tiplerinin sözdizimsel karmaşıklık üzerindeki kısa vadeli etkisi ile ilgili olarak doğrudan düzeltici geribildirim: T-birimi başına cümlecik T birimi başına koordinat cümleleri ve cümle başına cümlecik alanlarında kontrol grubundan istatistiksel olarak anlamlı ölçüde üstün gelmiştir. Doğrudan düzeltici geribildirim ayrıca, cümle başına cümlecik alanında metadilbilimsel açıklamadan istatistiksel olarak anlamlı ölçüde daha iyi bir performans göstermiştir. Metadilbilimsel açıklama, sözdizimsel karmaşıklık endeksleri bakımından kontrol grubundan anlamlı ölçüde farklı bulunmamıştır. Diğer sözdizimsel karmaşıklık endeksleri konusunda istatistiksel olarak anlamlı bir fark görülmemiştir. İki haftalık aradan sonra, gecikmeli son testin sonuçları doğrudan düzeltici geribildirim; T-birimi başına cümlecik ve cümle başına cümlecik alanlarında kontrol grubunu istatistiksel olarak anlamlı şekilde geride bıraktığını göstermiştir. Doğrudan düzeltici geribildirim ayrıca cümle başına cümlecikte metadilbilimsel açıklama 'yi istatistiksel olarak anlamlı şekilde geride bırakmıştır ancak doğrudan düzeltici geribildirim 'in T birimi başına koordinat cümleleri üzerindeki pozitif etkisi iki haftadan fazla sürmemiştir (yani, doğrudan düzeltici geribildirim, T birimi başına koordinat cümleleri alanında gecikmiş son test içinde kontrol gruptan anlamlı ölçüde farklı bulunmamıştır.). Uzun vadede, sözdizimsel karmaşıklık endeksleri bakımından kontrol gruptan önemli ölçüde farklı bulunmamıştır. Pedagojik ve teorik uygulamalar tartışılmış ve gelecekteki araştırmalar için öneriler verilmiştir.

Anahtar Kelimeler : kapsamlı (odaklanmamış) yazılı düzeltici geribildirim, doğrudan düzeltici geribildirim, metadilbilimsel açıklama, gramer bilgisi, sözdizimsel doğruluk, sözdizimsel karmaşıklık.

Sayfa Adedi : 189

Danışman : Doç. Dr. Hacer Hande UYSAL GÜRDAL

COMPREHENSIVE WRITTEN CORRECTIVE FEEDBACK: EFFECTS ON L2 SYNTACTIC ACCURACY AND COMPLEXITY IN WRITING

PhD Dissertation

Mohammadreza Valizadeh

GAZI UNIVERSITY

INSTITUTE OF EDUCATIONAL SCIENCES

July, 2019

ABSTRACT

This quasi-experimental study, which used a pretest-intervention-posttest-delayed posttest design, aimed at investigating the effects of two different types of comprehensive (unfocused) Written Corrective Feedback (WCF) strategies: direct corrective feedback (DCF), and metalinguistic explanation (ME) on L2 learners' both syntactic accuracy and syntactic complexity in writing in the short and long term (i.e., after a two-week interval). The participants were 90 Turkish EFL learners between the ages of 18 to 23, who enrolled in upper-intermediate essay writing classes of preparatory university courses. To control the L2 proficiency level of the participants and to ensure that they were homogeneous, first, they were given the pen-and-paper version of Quick Placement Test (QPT). To decide on the students' language level, Geranpayeh's (2003) guideline was used. Those who got scores ranging from 40 to 47 out of 60 (i.e., the upper-intermediate level), were selected for the IELTS writing Task 2 test, which was utilized to measure the syntactic accuracy and syntactic complexity in the written texts and ensure the homogeneity of the students in terms of their written accuracy and complexity. The students who were homogeneous in this regard were assigned to three groups, namely DCF (n = 30), ME (n = 30), who received feedback on their grammatical errors and a control group, namely 'No Feedback', abbreviated as NF (n = 30) who were provided with feedback only on content, orthography, and organization, but not on grammatical errors. The treatment/control period lasted for 5 weeks. Every week the participants in each group wrote an argumentative essay in class and turned it in; the teacher-researcher provided the specified feedback to each group. Then, the students were required to revise their corrected text. Next, on the 1st session of Week 6, the posttest was administered. No work on writing was done during the 6th and 7th weeks. Then, in the 8th week, the delayed posttest was implemented. In order to assess the syntactic accuracy, the following formula was used: [total number of syntactic

errors/total number of words] × 100. Besides, to calculate the syntactic complexity, Lu's (2010) web-based L2 Syntactic Complexity Analyzer, entitled, 'Web-based L2SCA: Single Mode' was utilized. In the present study, 10 out of 14 indices were analyzed. These ten syntactic complexity measures included: mean length of clause (MLC), mean length of sentence (MLS), mean length of the T-unit (MLT), clauses per sentence (C/S), clauses per T-unit (C/T), complex T-units per T-unit (CT/T), dependent clauses per clause (DC/C), coordinate phrases per T-unit (CP/T), complex nominals per T-unit (CN/T), and verb phrases per T-unit (VP/T). Eventually, both experimental groups significantly outperformed the control group in terms of the syntactic accuracy of the written texts both in the short and long term; and there was not a significant difference in the DCF and ME both in the short and long term. With regard to the short-term effect of the comprehensive WCF types on the syntactic complexity, the DCF statistically significantly outperformed the NF in the C/T, CP/T, and C/S of the texts. DCF also statistically significantly outperformed the ME in the C/S of the text. The ME was not significantly different from the NF in terms of the syntactic complexity indices. No statistically significant differences were seen regarding the other syntactic complexity indices. After the two-week interval, the results of the delayed posttest indicated the DCF still statistically significantly outperformed the NF in terms of the C/T and C/S. DCF also still statistically significantly outperformed the ME in C/S, but positive effect of DCF on CP/T was not durable over a two-week period (i.e., the DCF was not significantly different from the NF in terms of the CP/T in the delayed posttest). In the long-term, the ME was not significantly different from the NF in terms of the syntactic complexity indices. Pedagogical and theoretical implications are discussed, as well as suggestions for future research.

Key Words : comprehensive (unfocused) written corrective feedback, direct corrective feedback, metalinguistic explanation, grammatical knowledge, syntactic accuracy, syntactic complexity

Page Number : 189

Supervisor : Assoc. Prof. Dr. Hacer Hande Uysal Gürdal

TABLE OF CONTENTS

ÖZ v

ABSTRACT vii

LIST OF TABLES xiv

LIST OF FIGURES xvii

CHAPTER I 1

INTRODUCTION 1

1.1. Background 1

1.2. Statement of the problem 3

1.3. Purpose of the Study 4

1.4. Significance of the Study and Rationales for Selecting the Feedback Types 5

1.5. Research Questions 7

1.6. Delimitations 7

1.7. Definition of Key Terms 8

CHAPTER II 11

REVIEW OF RELATED LITERATURE 11

2.1. Introduction 11

2.2. Theoretical Perspectives on Error and WCF in SLA 11

2.2.1. Cognitive Perspective 11

2.2.1.1. Information Processing Models and Theories 12

2.2.1.2. The Interactionist Perspective 12

2.2.1.3. Noticing Hypothesis 13

2.2.1.4. Output Hypothesis 14

2.2.2. Socio-Cultural Perspective: A Different Interactionist Perspective 14

2.3. Perspectives of L2 Writing Theorists and Researchers on the Role of WCF in SLA 15

2.4. Research on WCF in L2 Writing 17

2.5. Studies on the Effect of Unfocused WCF on L2 Accuracy Gains 18

2.6. Studies Comparing the Effect of Focused with Unfocused WCF on Accuracy Gains 20

2.7. Studies on the Effect of Unfocused WCF on L2 Accuracy and Complexity Gains 22

2.8. Studies on the Effect of Direct WCF 24

2.9. Studies on the Effect of Metalinguistic CF on L2 Accuracy Gains 25

2.10. Studies Comparing the Effect of Direct CF with Metalinguistic CF on Accuracy Gains 25

2.11. Studies on the Potential Mediating Factors in WCF Efficacy 28

2.12. The Necessity to Do Further Studies 30

CHAPTER III 31

METHODOLOGY 31

3.1. Introduction 31

3.2. Research Questions 31

3.3. Research Setting 31

3.4. Research Participants 32

3.5. Research Design 32

3.5.1. Variables in the Study 33

3.6. Instruments 33

3.6.1. Students' Biodata Information 34

3.6.2. Quick Placement Test 34

3.6.2.1. Quick Placement Test Reliability and Validity 34

3.6.3. A Pretest, a Posttest, a Delayed-posttest, and Classroom Tasks of Writing 35

3.6.4. Syntactic Accuracy Measure 35

3.6.5. Syntactic Complexity Measure 36

3.7. Data Collection Procedure 37

3.8. Data Analysis 38

CHAPTER IV 41

RESULTS AND DISCUSSIONS 41

4.1. Introduction 41

4.2. Inter-rater Reliability Indices 41

4.3. The Normality Tests 41

4.3.1. Numerical Methods of Assessing Normality for QPT and Syntactic Accuracy 42

4.3.2. Numerical Methods of Assessing Normality for Syntactic Complexity 42

4.4. Ensuring the Homogeneity of the Groups 43

4.4.1. Grammatical Knowledge 43

4.4.2. Syntactic Accuracy 44

4.4.3. Syntactic Complexity 45

4.5. Answer to the First Research Question 48

4.5.1. The Short-term Effect: Post-test Result 48

4.5.2. The Long-term Effect: Delayed-post-Test Result 49

4.6. Answer to the Second Research Question 51

4.6.1. The Short-term Effect: Posttest Result 51

4.6.2. The Long-term Effect: Delayed-post-Test Result 52

4.7. Answer to The Third Research Question 53

4.7.1. The Short-term Effect: Post-test Result 53

4.7.2. The Long-term Effect: Delayed-post-test Result 54

4.8. Answer to the Fourth Research Question 55

4.8.1. The Short-term Effect: Post-test Result 55

4.8.2. The Long-term Effect: Delayed-post-test Result 64

4.9. Discussion 73

4.9.1. Syntactic Accuracy 73

4.9.2. Syntactic Complexity 75

CHAPTER V 79

CONCLUSIONS 79

5.1. Overview 79

5.2. Summary and Conclusion 79

5.3. Theoretical Implications 80

5.4. Pedagogical Implications 81

5.5. Limitations of the Study 83

5.6. Suggestions for Future Research 83

REFERENCES 87

APPENDİXES 109

Appendix 1: Biodata Information 110

Appendix 2: Quick Placement Test (QPT) 111

Appendix 3: Writing Topics 117

Appendix 4: Histograms for QPT 119

Appendix 5: Histograms for Syntactic Accuracy Measures 121

Appendix 6: Histograms for Pretest of Syntactic Complexity 126

Appendix 7: Histograms for Posttest of Syntactic Complexity 141

Appendix 8: Histograms for Delayed-pretest of Syntactic Complexity 156

Appendix 9: Normality Tests for QPT and Syntactic Accuracy 171

Appendix 10: Normality Tests for Pretest of Syntactic Complexity 172

Appendix 11: Normality Tests for Posttest and Delayed-posttest of Syntactic Complexity 174

Appendix 12: Correlations Between Dependent Variables for Pretest of Syntactic Complexity 178

Appendix 13: Scatterplot for Linearity Assumption for Pretest of Syntactic Complexity 180

Appendix 14: Correlations Between Dependent Variables for Posttest of Syntactic Complexity 182

Appendix 15: Scatterplot for Linearity Assumption for Posttest of Syntactic Complexity 184

Appendix 16: Correlations Between Dependent Variables for Delayed-posttest of Syntactic Complexity 186

Appendix 17: Scatterplot for Linearity Assumption for Delayed-posttest of Syntactic Complexity 188

LIST OF TABLES

Table 3.1. *Ten Syntactic Complexity Measures* 36

Table 4.1. *Inter-Rater Reliability Indices for Syntactic Accuracy* 41

Table 4.2. *Descriptives for QPT* 43

Table 4.3. *Levene's Test for Homogeneity of Variances for QPT* 43

Table 4.4. *ANOVA for the QPT* 44

Table 4.5. *Descriptives for Pretest of Syntactic Accuracy* 44

Table 4.6. *Levene's Test for Homogeneity of Variances for Pretest of Syntactic Accuracy* 44

Table 4.7. *ANOVA for Pretest of Syntactic Accuracy* 45

Table 4.8. *Box's Test of Equality of Covariance Matrices for Pretest of Syntactic Complexity* 46

Table 4.9. *Levene's Test of Equality of Error Variances for Pretest of Syntactic Complexity* 47

Table 4.10. *Multivariate Tests for Pretest of Syntactic Complexity* 47

Table 4.11. *Levene's Test for Homogeneity of Variances for Pretest of CP/T* 48

Table 4.12. *ANOVA for Pretest of CP/T* 48

Table 4.13. *Descriptive Statistics Posttest of Syntactic Accuracy by DCF & NF* 49

Table 4.14. *Independent T-test Posttest of Syntactic Accuracy by DCF & NF* 49

Table 4.15. *Descriptive Statistics Delayed-Posttest of Syntactic Accuracy by DCF & NF* 50

Table 4.16. *Independent T-test Delayed-Posttest of Syntactic Accuracy by DCF & NF* 50

Table 4.17. *Descriptive Statistics Posttest of Syntactic Accuracy by ME & NF* 51

Table 4.18. *Independent T-test Posttest of Syntactic Accuracy by ME & NF* 51

Table 4.19. *Descriptive Statistics Delayed-Posttest of Syntactic Accuracy by ME & NF*... 52

Table 4.20. *Independent T-test Delayed-Posttest of Syntactic Accuracy by ME & NF*..... 52

Table 4.21. *Descriptive Statistics Posttest of Syntactic Accuracy by DCF & ME* 53

Table 4.22. *Independent T-test Posttest of Syntactic Accuracy by DCF & ME* 54

Table 4.23. *Descriptive Statistics Delayed-Posttest of Syntactic Accuracy by DCF and ME*.

..... 54

Table 4.24. *Independent T-test Delayed-Posttest of Syntactic Accuracy by DCF and ME*

..... 55

Table 4.25. *Box's Test of Equality of Covariance Matrices for Posttest of Syntactic Complexity* 57

Table 4.26. *Levene's Test of Equality of Error Variances for Posttest of Syntactic Complexity* 57

Table 4.27. *Multivariate Tests for Posttest of Syntactic Complexity*..... 58

Table 4.28. *Tests of Between-Subjects Effects for Posttest of Syntactic Complexity*..... 59

Table 4.29. *Tukey HSD: Multiple Comparisons for Posttest of Syntactic Complexity for C/T*..... 60

Table 4.30. *Descriptives for Posttest of C/S* 61

Table 4.31. *Levene's Test for Homogeneity of Variances for posttest of C/S* 62

Table 4.32. *ANOVA for Posttest of C/S* 62

Table 4.33. *Tukey HSD: Multiple Comparisons for Posttest of C/S*..... 62

Table 4.34. *Descriptives for Posttest of CP/T*..... 63

Table 4.35. *Levene's Test for Homogeneity of Variances for Posttest of CP/T* 63

Table 4.36. *ANOVA for Posttest of CP/T*..... 63

Table 4.37. *Tukey HSD: Multiple Comparisons for Posttest of CP/T*..... 64

Table 4.38. *Box's Test of Equality of Covariance Matrices for Delayed-Posttest of Syntactic Complexity*..... 65

Table 4.39. *Levene's Test of Equality of Error Variances for Delayed-Posttest of Syntactic Complexity* 66

Table 4.40. *Multivariate Tests for Delayed-Posttest of Syntactic Complexity* 67

Table 4.41. *Tests of Between-Subjects Effects for Delayed-Posttest of Syntactic Complexity* 68

Table 4.42. *Tukey HSD: Multiple Comparisons for Delayed-Posttest of Syntactic Complexity for C/T* 69

Table 4.43. *Descriptives for Delayed-Posttest of C/S* 70

Table 4.44. *Levene's Test for Homogeneity of Variances for Delayed-Posttest of C/S* 71

Table 4.45. *ANOVA for Delayed-Posttest of C/S* 71

Table 4.46. Tukey *HSD: Multiple Comparisons for Delayed-Posttest of C/S* 71

Table 4.47. *Descriptives for Delayed-Posttest of CP/T* 72

Table 4.48. *Levene's Test for Homogeneity of Variances for Delayed Posttest of CP/T* 72

Table 4.49. *Robust Tests of Equality of Means for Delayed-Posttest of CP/T* 72

LIST OF FIGURES

Figure 4.1. Estimated Marginal Means for Posttest of C/T 61

Figure 4.2. Estimated Marginal Means for Delayed-Posttest of C/T 70

CHAPTER I

INTRODUCTION

1.1. Background

It is well acknowledged that output production is a valuable source for the acquisition of second language (L2) forms, as it pushes learners to engage in deeper processing of language and with more mental effort than does the language input (i.e., listening and reading) (e.g., Ellis, 2003, 2005; Skehan, 1998). Output production allows learners to reflect on their metalinguistic knowledge and test hypotheses about the L2 grammar (Swain, 1985, 1995). Output can also promote noticing, which is an essential condition for L2 acquisition (Schmidt, 1990; 1993) as it helps learners become aware of the gaps and problems in their interlanguage system, solve their grammatical problems, and develop their interlanguage competence (Izumi, 2002; Swain, 2005). These are specified as the metalinguistic, hypothesis testing, and noticing functions of the output (Swain, 1995).

The ability to write in an L2, as a form of output production, in the 21st century, has become more important than ever because of globalization requiring written communication across languages and cultures and in many fields (Weigle, 2014). Given such necessity and importance, writing is a complex process due to the fact that it requires the skillful coordination of cognitive, linguistic, and social processes and resources (Bhatia, 2002; Hayes, 1996; Kellogg, 1996). Focusing on planning and organizing ideas as well as spelling, punctuation and word choice can be a daunting task (Richards & Renandya, 2002). Therefore, one of the challenging tasks second language (henceforth, L2) teachers face, is guiding their students through the difficult process of becoming skilled writers in the L2 (Farah, 2014; Van Beuningen, 2011; Weigle, 2014). Further, whereas writing in an L2 can be cognitively challenging, it can also at the same time be of vital importance to L2 development (Van Beuningen, 2011; Williams, 2012).

Nevertheless, Swain (1991) noted that output by itself cannot necessarily accomplish all these functions; therefore, some kind of corrective feedback (CF) is also needed during the output production for L2 linguistic development. Swain stated "if students are given insufficient feedback or no feedback regarding the extent to which their messages have successfully (accurately, appropriately, and coherently) been conveyed, output may not serve these roles" (Swain, 1991, p. 98). Similarly, Han (2002) and Havranek (2002) stated learners' output had better be accompanied by CF in order to be beneficial for the language learning process, because L2 instruction without any attention to grammar or error correction is likely to result in fossilization of errors and use of poor L2 grammar in both oral and written production (Doughty & Williams, 1998; Schmidt, 1990; Swain, 1985; White, 1987). Furthermore, grammatical forms, especially as opposed to learners' first language, that are rare in the input, and that are not regular, cannot be learned simply through exposure alone, but some intervention or CF is necessary on these forms (Larsen-freeman, 2003; R. Sheen, 2003). As a matter of fact, adult L2 learning requires some degree of negative input (VanPatten, 2015) in the form of CF (Ellis & Shintani, 2014) to guide learners to notice the gap between the ideal and their existing knowledge, recognize what is lacking or wrong in their production and modify their interlanguage (Susan M. Gass & Varonis, 2004; R. W. Schmidt, 1995; 2001), possibly as a result of the established explicit knowledge (DeKeyser, 2015).

The most commonly used feedback type, targeting linguistic problems, is error correction or CF, which can be supplied on L2 learners' writing in several ways, regarding "their explicitness, their focus, the person providing the feedback, the feedback medium", and so on (Van Beuningen, 2011, p. 3). In addition, previous studies have indicated that L2 learners have positive attitudes towards teacher's provision of written corrective feedback (henceforth, WCF) and expect their writing teachers to provide it (Alkhatib, 2015; Chandler, 2003; Cohen & Cavalcanti, 1990; Cumming, 1995; Diab, 2005b, 2005a; Ferris, 1995b, 2004; Ferris & Roberts, 2001; Grami, 2005; Hajian, Farahani, & Shirazi, 2014; Hamouda, 2011; Hedgcock & Lefkowitz, 1994, 1996; Horbacauskiene & Kasperaviciene, 2015; F. Hyland, 1998; Lee, 1997, 2004; Leki, 1991; Mahfoodh & Pandian, 2011; Zacharias, 2007). However, Truscott (1996) claimed that CF was not ffective and is even potentially harmful. In light of these positive suggestions and negative points on CF on language production, WCF remains a controversial, but important issue to be explored in L2 writing classrooms (Brown, 2017; Ferris, 2014).

1.2. Statement of the Problem

Truscott (1996) argued CF is not only ineffective but it can even be harmful; thus, providing it must be stopped. Truscott referred to syntactic knowledge and argued that WCF cannot facilitate the acquisition of syntactic knowledge. Moreover, Truscott (2004, 2007) speculated that if learners were faced with corrective feedback (henceforth, CF), it is highly probable that they do not use complex structures which they were not confident whether or not they are able to use accurately. From a theoretical perspective, limited capacity models of attention (Skehan, 1998) supports what Truscott stated because according to these models, learners may concentrate their efforts on gaining greater control over items which had already been internalized and are more stable (interlanguage) and avoid expanding their L2 knowledge system (Skehan & Foster, 2001). Since then, Truscott's argument has yielded several counter responses from a number of researchers (Bruton, 2009a, 2010; Ferris, 1999, 2004, 2006; Lyster, Lightbown, & Spada, 1999). Truscott (2007), despite his negative attitudes towards feedback, claimed that WCF is only likely to have a positive effect in helping learners in treating errors that "are relatively simple and can be treated as discrete items" (p. 258).

Researchers have investigated this issue with various methodological approaches (Ellis, 2010; Ferris, 1999; Guénette, 2007; Liu & Brown, 2015; Storch, 2010) and explored the effect of different WCF strategies on L2 learners' accuracy in writing (Benson & DeKeyser, 2018; Bitchener, 2008; Bitchener, East, & Cartner, 2010; Bitchener, Young, & Cameron, 2005; Bitchener & Knoch, 2008; Bonilla López, Van Steendam, Speelman, & Buyse, 2018; Ellis, Sheen, Murakami, & Takashima, 2008; Ferris, 2006; Ferris & Roberts, 2001; Frear & Chiu, 2015; Karim & Nassaji, 2018; Liu, 2016; Mawlawi Diab, 2015; Nassaji, 2011; Sadeghi, Abolfazli Khonb, & Gheitaranzadeh, 2013; Santos, López-Serrano, & Manchón, 2010; Sato, 2012; Y. Sheen, 2007, 2010b; Y. Sheen, Wright, & Moldawa, 2009; Truscott, 2007; Van Beuningen, De Jong, & Kuiken, 2012). Therefore, feedback related "studies using measures of accuracy are frequent in the second language writing literature" (Polio & Shea, 2014, p. 10). However, the majority of them explored the effects of focused or selective feedback on EFL writing (Benson & DeKeyser, 2018; Bitchener, 2008; Bitchener et al., 2010, 2005; Bitchener & Knoch, 2008, 2009, 2010b; Y. Sheen, 2007, 2010a; Y. Sheen et al., 2009), whereas "evidence on the language learning potential of unfocused CF, which involves comprehensive correction of every error in students' writing, is scarce" (Van Beuningen et al., 2012, p. 5).

Furthermore, although many studies focused on the effects of WCF on accuracy, to date, only few studies have investigated the effect of WCF strategies on developing syntactic complexity in L2 writing (Chandler, 2003; Fazilatfar, Fallah, Hamavandi, & Rostamian, 2014; Robb, Ross, & Shortreed, 1986; Sheppard, 1992; Van Beuningen et al., 2012).

These few studies on the effect of WCF on written complexity have revealed conflicting results (Bitchener & Ferris, 2012). For example, Robb et al., (1986), who studied indirect corrective feedback on the complexity, found a significant positive effect, but later, Sheppard (1992) found a negative effect for WCF on structural complexity. On the other hand, Chandler (2003) concluded that WCF did not have any effect on the complexity of L2 learner writing. Most recently, Van Beuningen et al., (2012) also reported that WCF did not result in writing simplified texts in their posttests. Therefore, considering that the results of WCF studies on written syntactic complexity are inconclusive, this issue needs more meticulous investigation.

Furthermore, several scholars have focused researchers' attention on individual variables that could influence how and when WCF can be the most effective (Chen, Nassaji, & Liu, 2016; Cohen, 2012; Evans, Hartshorn, McCollum, & Wolfersberger, 2010; Ferris, 2004, 2010; Guénette, 2007; F. Hyland, 1998, 2000; K. Hyland & Hyland, 2006; Kang & Han, 2015; Kormos, 2012; Y. Sheen, 2011; Storch, 2010). For instance, as Guénette (2007), Storch (2010), and Van Beuningen (2010) argued, the success or failure of WCF will depend on a collection of variables including the students' proficiency level, one of the learner-internal factors which can intervene the effect of WCF; such mediating variables must be taken into consideration. As a result, Kang and Han (2015) concluded that "much is still unknown about which types of corrective feedback are more effective." (p. 12).

1.3. Purpose of the Study

First, this study aimed at investigating the effects of two different types of WCF strategies: direct corrective feedback (henceforth, DCF), and metalinguistic explanation (henceforth, ME) on L2 learners' both syntactic accuracy and syntactic complexity in writing. DCF consists of an indication of the error and provision of the corresponding correct L2 form (Ellis, 2009); however, by ME, the learners are not provided with the correct form, but "the teacher numbers errors in text and writes a grammatical description for each numbered error at the bottom of the text" (Ellis, 2009, p. 98) or on a separate paper, which are attached to the student's written text. Syntactic accuracy means producing syntactic error-

free texts, while syntactic complexity means what Ortega (2015) defined as "the sophistication, variety, diversity, or elaboratedness of grammatical resources exhibited in language production" (p. 86). This study, further, sought whether any effect of WCF on accuracy and complexity gains would last in the long term.

In addition, students' L2 proficiency and writing levels were controlled. As Frodesen and Holten (2003, as cited in Guénette, 2007) argued, to be able to compare the efficacy of different WCF strategies, the students' English proficiency level had to be carefully controlled. Therefore, this study, first, explored the English proficiency level of the students via Oxford Quick Placement Test and formed the groups who were homogenous to prevent any interference of grammatical knowledge factor into the study. Then, a sample of the Task 2 of the General Writing Module of IELTS was administered as the pretest and statistically analyzed to ensure that the participants were homogeneous regarding their written syntactic accuracy and complexity.

1.4. Significance of the Study and Rationale for Selecting the Feedback Types

Many L2 learners and those getting prepared for the demands of academic writing need to improve not only their written accuracy but also the syntactic complexity in L2 writing (Balanga et al., 2016; Brown, 2017). On the other hand, "language teachers are equally committed to helping their learners develop fluency in their writing and, especially at higher proficiency levels, greater levels of complexity in their writing" (Bitchener & Ferris, 2012, p. 179).

Because this study aimed at investigating the syntactic complexity of the written texts, the unfocused feedback would be the most appropriate one as it was also used in the previous studies on the syntactic complexity (Chandler, 2003; Fazilatfar et al., 2014; Robb et al., 1986; Sheppard, 1992; Van Beuningen et al., 2012). Besides, in focused feedback studies, tasks should be specifically designed to allow the pre-selection of one or two grammatical structures to be focused and to find out whether or not the treatment is effective on correcting these errors (Benson & DeKeyser, 2018; Bitchener, 2008; Bitchener & Knoch, 2008, 2009, 2010b; Y. Sheen et al., 2009; Shintani & Ellis, 2013). Natural or free-writing tasks cannot be used for this purpose as was pinpointed by Shintani and Ellis (2015). However, while controlling the task type to ensure that the participants used the focused target structures, exploration of the effect of the treatment on syntactic complexity would

not be achieved. Therefore, in the present study, free-writing tasks were used, so that assessing the syntactic complexity of the written texts could be feasible.

With regard to the effect of unfocused feedback on accuracy, because the unfocused feedback is the most widely used type of feedback by teachers (Ferris, 2006; Guénette, 2012; Lee, 2004, 2008; Van Beuningen, 2010), it is more ecologically valid than the focused feedback (Ferris, 2010). The unfocused or comprehensive feedback is also concerned with the writing as a whole instead of the writing as a way to practice grammar (Bruton, 2009b; Van Beuningen, 2010), which was the aim in the current study. Therefore, the comprehensive (unfocused) feedback serves the need for more authentic CF methodologies, which focus "on the accurate production of all aspects of writing, simultaneously" (Hartshorn et al., 2010, p. 89). In addition, in case the learners commit a range of written errors, "a limited CF focus does not address the need to individualize feedback according to students' different strengths and weaknesses" (Ferris, 2010, p. 192). Moreover, studies on unfocused WCF have been reported to be scarce (Van Beuningen et al., 2012). Van Beuningen, et al., (2012) argued that "it is important to produce more evidence about the singular contribution of unfocused or comprehensive CF on the accuracy of new pieces of writing" (p. 6). Additionally, as Bitchener and Ferris (2012) stated, due to the fact that the results available about the effectiveness of providing learners with unfocused WCF are conflicting, "further research is required on the relative benefits of unfocused feedback" (p. 57). Further, as the participants in this study were at upper-intermediate level, comprehensive WCF could be more useful for them (Bitchener & Ferris, 2012).

Among the WCF strategies, DCF and ME were selected to be compared with each other as well as with 'no feedback' condition. Because DCF consists of an indication of the error and provision of the corresponding correct L2 form (Ellis, 2009), the learners are provided with adequate information to fix complex linguistic errors, such as the syntactic ones; therefore, the learners will be able to instantly internalize the correct form (Chandler, 2003). Additionally, DCF is more likely to facilitate learning in case the learners' level of (meta)linguistic competence is not high (i.e., they "have no or ill-formed explicit or implicit knowledge of a grammatical feature") to be able to self-correct their errors (Shintani & Ellis, 2013, p. 288). Because in this study, the unfocused/comprehensive feedback was selected and all the learners' syntactic errors were corrected, the DCF was used to maximize the potential effect of the CF.

Regarding the ME, although the comprehensive or unfocused feedback was provided and also the learners were not provided with the correct form, they were provided with explicit comments on their errors by provision of the grammatical description (Ellis, 2009) as it is argued that the ME causes the students to engage in deeper level of processing due to the fact that they need to utilize the ME to correct their own errors (Shintani & Ellis, 2013). Because the participants in the present study were the upper-intermediate ones, it was assumed that they were able to use the ME to identify and self-correct their own errors.

Consequently, the present research, which was motivated by the above-mentioned issues, hoped to shed light on these issues for L2 teachers.

1.5. Research Questions

The study addressed the following research questions:

1) Is there any significant difference between the effect of DCF and NF (i.e., no feedback) on developing learners' syntactic accuracy in writing in the short and long term?
2) Is there any significant difference between the effect of ME and NF on developing learners' syntactic accuracy in writing in the short and long term?
3) Is there any significant difference between the effect of DCF and ME on developing learners' syntactic accuracy in writing in the short and long term?
4) To what extent are the comprehensive DCF and ME effective on developing EFL learners' written syntactic complexity, relative to each other and to NF (i.e., no feedback) in the short and long term?

1.6. Delimitations

Delimitations are the elements over which the researcher has some control. They are characteristics selected by the researcher to describe the scope and define the boundaries of the study Dusick (2011). The present study started with the following delimitations: (1) the researcher decided to work with learners at the upper-intermediate level of English proficiency; (2) the research was conducted in prep-university courses focusing the general English with students studying all the skills, not just writing; thus, the results may not necessarily be generalizable to the context of university advanced composition classes and students who attend them. And (3) the non-random convenience sampling method, which was used in this study, (i.e. the researcher use intact classes) can compromise the external validity of the study.

1.7. Definition of Key Terms

Writing: As Sokolic (2003) explained, writing can be defined by a series of contrasts:

- It is not only a *physical* but also a *mental* act. Writing is a mental work because writers think to invent ideas and think about how to express them. It is also a physical act because words are used to organize ideas, and then ideas are organized into statements and paragraphs that will be clear to a reader.
- Writing's purpose is both to *express* and *impress*. Writers express their *own* ideas or feelings, but they express them in a way to be intelligible to readers or audience.
- Writing is both a *process* and a *product*. The process of writing includes thinking, organizing, drafting and editing a written piece. What the audience ultimately sees is a product – a piece of writing.

Feedback: "In teaching, feedback refers to comments or other information that learners receive concerning their success on learning tasks or tests, either from the teacher or other persons." (Richards & Schmidt, 2010, p. 217).

Error Correction: Richards and Schmidt (2010) defined error correction as strategies used by a teacher or more advanced learner to correct errors in a learner's production. Error correction may be direct (teacher supplies the correct form) or indirect (the teacher points out the problem and asks the learner to correct it if possible). (p. 202)

Direct Corrective Feedback: Direct corrective feedback is one of the strategies for providing CF, through which "the teacher provides the student with the correct form." (Ellis, 2009, p. 98).

Metalinguistic Explanation: Metalinguistic explanation is one of the strategies for providing CF, through which "teacher numbers errors in text and writes a grammatical description for each numbered error at the bottom of the text." (Ellis, 2009, p. 98).

The Focus of the Feedback: The focus of feedback "concerns whether the teacher attempts to correct all (or most) of the students' errors or selects one or two specific types of errors to correct." (Ellis, 2009, p. 98).

Unfocused CF: The unfocused corrective feedback is extensive (i.e., the teacher attempts to correct all (or most) of the students' errors) (Ellis, 2009).

Accuracy: Accuracy "refers to the ability to produce grammatically correct sentences" (Richards & Schmidt, 2010, p. 203).

Syntactic Complexity: Syntactic complexity is defined as "the sophistication, variety, diversity, or elaboratedness of grammatical resources exhibited in language production" (Ortega, 2015, p. 86).

Language Proficiency: Language proficiency is defined by Richards and Schmidt (2010) as the degree of skill with which a person can use a language, such as how well a person can read, write, speak, or understand language. This can be contrasted with language achievement, which describes language ability as a result of learning. Proficiency may be measured through the use of a proficiency test. (p. 321)

Proficiency Level: The proficiency level is "a description of a language learner's level of performance in a target language, often described in terms of *beginner level*, *intermediate level* or *advanced level*." (Richards & Schmidt, 2010, p. 464).

Essay: Essay (in composition) is "a longer piece of writing, particularly one that is written by a student as part of a course of study or by a writer writing for publication which expresses the writer's viewpoint on a topic." (Richards & Schmidt, 2010, p. 203).

Argumentative writing: Based on Richards and Schmidt (2010), Argumentative writing attempts to support a controversial point or defend a position on which there is a difference of opinion. ESL writing programmes have often been based on the assumption that novice writers should begin with the simplest mode – the descriptive essay – and gradually move to learning the most difficult – the argumentative one. (p. 372)

CHAPTER 2

REVIEW OF LITERATURE

2.1. Introduction

This chapter presents a review of the related literature in the area of the current study. It begins with some of the main theoretical perspectives (cognitive and socio-cultural ones) on error and WCF in SLA. Additionally, it focuses on perspectives of L2 writing theorists and researchers on the role of WCF in SLA. The chapter continues reporting and explaining the relevant previous studies which have been conducted. Next, the potential factors which might mediate or mitigate the efficacy of the WCF will be discussed. Finally, the open issues for the present work will be explained.

2.2. Theoretical Perspectives on Error and WCF in SLA

The reason for expecting that error correction can be beneficial to language learning depends on various theoretical foundations (Van Beuningen, 2011), some of which will be mentioned in this section.

2.2.1. Cognitive Perspective.

According to the cognitive perspective, learning an L2 is a process in which learners develop an interlanguage system based on their experience with L2 input (Z. H. Han, 2002). Input refers to all the language a learner is exposed to, through either oral medium or visual one) (Gass & Mackey, 2015; Mackey, 2012; Richards & Schmidt, 2010). Nevertheless, the available language input is not the only essential ingredient for L2 acquisition (Mackey, 2012); the L2 input has to be noticed and mentally processed by the learner, so that it will lead to L2 development (i.e., turn into intake) (Gass & Mackey, 2015).

2.2.1.1. Information Processing Models and Theories.

According to McLaughlin's information processing model, learning occurs when, through repeated practice, declarative knowledge becomes automatic. In addition to practice, it is also hypothesized that learners make changes in their language systems via a process called 'restructuring'. Through the cognitive process of 'restructuring', the previously learned information that has been stored in separate categories is integrated, which promotes the learner's competence (McLaughlin, 1990). Anderson's ACT (Adaptive Control of Thought) model (Anderson, 1993) is similar to McLaughlin's model because it centers on the belief that practice results in automatization (Bitchener & Ferris, 2012). These models incorporate "the view that information can be processed in either a controlled or automatic manner and that learning includes a shift from controlled toward automatic processing" (Bitchener & Ferris, 2012, p. 12). It explains that intentional learning, for instance, via explicit instruction and corrective feedback, can play an important "role in the controlled phase and through 'practice' or 'repeated activation', language over time becomes automatized" (i.e., declarative knowledge becomes automatized procedural knowledge) (Bitchener & Ferris, 2012, p. 13). This theoretical view has also been supported by researchers (e.g., DeKeyser, 1997, 2001; Hulstijn, 1995; Schmidt, 1990, 1994, 1995; Swain, 1985; Swain & Lapkin, 1995). However, as Pienemann explained in his teachability hypothesis and processability theory (PT) (Pienemann, 1987, 1989, 1998), "information processing is unlikely to occur if the targeted linguistic forms and structures lie outside a learner's stage of 'readiness'" (Bitchener & Ferris, 2012, p. 15).

2.2.1.2. The Interactionist Perspective.

Some theorists who work mainly within an SLA framework think that language learning largely takes place through social interaction (Spada & Lightbown, 2010). As a matter of fact, it is now commonly accepted within the SLA literature that there is a strong connection between interaction and learning (Gass & Mackey, 2015; Mackey, 2012). Although the interactionist perspectives on SLA have focused mainly on the role of oral interaction between learners and their interlocutors, it does not mean that proposals and empirical findings emerging from this context are not relevant to the written context and the role of WCF in SLA (Bitchener & Ferris, 2012). According to the interaction approach, learning occurs when the learner is exposed to language, produces language and receives feedback on that production (Gass & Mackey, 2015). Thus, the case for including negative

evidence (including CF) as well as positive evidence as input for language learning has been promoted by the interactionist theorists (Bitchener & Ferris, 2012). For example, Long's (1996) Interaction Hypothesis proposed that

> environmental contributions to acquisition are mediated by selective attention and the learner's developing L2 processing capacity, and that these resources are brought together most usefully, although not exclusively, during negotiation of meaning. Negative feedback obtained during negotiation work or elsewhere may be facilitative of L2 development, at least for vocabulary, morphology and language-specific syntax, and essential for learning certain specifiable L1-L2 contrasts (p. 414).

Furthermore, Long's focus-on-form approach as a pedagogical intervention is supported in the SLA field (Long, 1996; Long & Robinson, 1998) and one of the identified pedagogical focus-on-form instrument is error correction (Ellis, 2005). CF is a personalized individualized reactive focus-on-form technique, drawing students' attention to form while doing a task. Learners only have to deal with WCF after meaning has been communicated while oral feedback will inevitably interrupt the communicative flow (Polio, Fleck, & Leder, 1998).

Therefore, based on the above-mentioned points as well as the quotation from Long, it is necessary to provide learners with negative evidence (including CF) as well as positive evidence, but a learner's processing capacity and degree of attention to "linguistic form may determine the extent to which L2 input becomes L2 intake" (Bitchener & Ferris, 2012, p. 17). Some researchers, such as Smith (1981, 1993) and Schmidt, (1990, 2001) also explained that in order for some feature of language to be acquired, it is not enough for the learner to be exposed to it through comprehensible input, rather it is essential for learners to pay attention to language form and structure for acquisition to occur. Without attention, in Doughty's (2003) opinion, learners' accuracy development could be expected to be slower, more difficult, and less successful.

2.2.1.3. Noticing Hypothesis.

Considering the above-mentioned information, one of the fundamental reasons of the focus-on-form technique can be seen in Schmidt's Noticing Hypothesis (Schmidt, 1990, 2001). Attention makes learners aware that there is "a mismatch or gap between what they can produce and what they need to produce, as well as between what they produce and

what target language speakers produce" (Schmidt, 2001, p. 6). This concept has been commonly referred to as noticing the gap (e.g. Schmidt and Frota, 1986, as cited in Van Beuningen, 2011). Ellis (1995) used the term *cognitive comparison* instead of *noticing the gap* because the learners must also notice even when the output and the input are the same.

Thus, conscious attention to linguistic form facilitates or even is required for interlanguage development, so it is expected that focus-on-form treaments such as CF can support the SLA process (DeKeyser, 1994; Han, 2002). In increasing learners' awareness of certain linguistic features, CF makes learners able to notice the gaps between their own interlanguage output and the target language input provided via feedback (Hulstijn & Schmidt, 1994). Consequently, these noticing actions could cause the learners' developing interlanguage grammar to be restructured (Gass, 1997; Long, 1996). Moreover, unlike oral feedback, when WCF is provided, learners have enough time – and therefore cognitive resources – to compare their output with the received CF, which raises the likelihood of learners noticing the gaps in their interlanguage (Polio et al., 1998; Y. Sheen, 2010b).

2.2.1.4. Output Hypothesis.

According to Swain's (1985, 1995) *output hypothesis*, learners, while producing output, are pushed to process language more deeply and with more mental effort than the language process during the receptive skills. Learner output, as a valuable source for acquisition, is also acknowledged by several scholars (e.g., Ellis, 2003, 2005; Skehan, 1998). Producing output allows learners to reflect on their metalinguistic knowledge and test hypotheses about the L2 grammar, so the produced output can promote noticing to be aware of the gaps and problems in their interlanguage system (Swain, 1985, 1995). Nevertheless, Swain (1991) added that output by itself does not necessarily do all these functions. Swain recognized the importance of CF by stating that "if students are given insufficient feedback or no feedback regarding the extent to which their messages have successfully (accurately, appropriately, and coherently) been conveyed, output may not serve these roles" (Swain, 1991, p. 98). Similarly, Han (2002) stated that learners' output had better be accompanied by CF so that it could be beneficial to the language learning process.

2.2.2. Socio-Cultural Perspective: A Different Interactionist Perspective.

Socio-cultural theory of human mental processing, based on Vygotsky, assumes that language development, as an example of cognitive development, occurs via social

inteactions; therefore, when learners collaborate and interact with native speakers of the L2 or teachers and more advanced more knowledgeable learners, their language abilities develop (Bitchener & Ferris, 2012) because they receive proper scaffolding; in other words, when learners receive help from teachers or more advanced learners (i.e., the help of other regulation), they will be able to use the L2 autonomously (i.e., be self-regulated) (Lantolf & Thorne, 2007; Lantolf, Thorne, & Poehner, 2015; Mackey, 2012). In fact, when learners are not yet able to use the L2 autonomously as procedural knowledge, but their L2 ability levels can be raised with the help of other regulation and scaffolding, it is said that the learners are in their zone of proximal development (ZPD) (Bitchener & Ferris, 2012). Empirical research into the L2 development in the ZPD has also supported this theory (Aljaafreh & Lantolf, 1994; Nassaji & Swain, 2000).

Teachers or more advanced learners (i.e., other regulators) may utilize various strategies such as repetition, simplification, modeling, and corrective feedback (Lantolf & Thorne, 2007). Aljaafreh and Lantolf (1994), in a longitudinal study of adult ESL learners, analyzed the relationship between the one-to-one teacher WCF (other regulator) and language learning. Aljaafreh and Lantolf (1994) utilized a "regulatory scale" (scaffolding) to indicate how the instructor's treatment could be ranged on a continuum from explicit to implicit CF. It was found that the more implicit the type of CF was, the closer the student got toward more independent and self-regulated performance, which was considerd as positive evidence of L2 learning and development.

Then, Nassaji and Swain (2000) utilized the regulatory scale to scaffold properly in their learners' ZPD to formally test whether effective scaffolding is contingent on the condition of the two Korean ESL learner's ZPD. One of the learners received randomly selected CF and the other one received negotiated ZPD-related CF. It was found that the randomly selected CF was not helpful in moving the learner toward self-regulation whereas the negotiated ZPD-related CF was.

2.3. Perspectives of L2 Writing Theorists and Researchers on the Role of WCF in SLA

Truscott (1996) expressed his opposition towards providing the error correction on students' written texts. He proposed that error correction had better be abandoned, and he provided some arguments for what he proposed. Truscott believed that because the learning processes facilitating the development of a learner's interlanguage is very

complex, provision of CF by teachers, which is a simple transfer of information, cannot enable their learners to correct their grammatical errors. Based on Truscott, due to the fact that to learn syntax, morphology, and lexis, the learner needs to understand not only form but also meaning and use, no individual form of CF can help the knowledge of all linguistic forms and structures be acquired.

Truscott's (1996) second argument against the WCF provision, which is more of a pedagogical reason than a theoretical one, was that the L2 teachers are not capable of providing WCF at a time when the learner is cognitively 'ready' to learn a special form or structure. Truscott also proposed that WCF cannot be effective if it covers a too comprehensive range of error categories.

Truscott's (1996) third argument against the provision of WCF is that learning that results from receving the WCF is probably only "pseudo-learning". He elaborated on the term as "a superficial and possibly transient form of knowledge" (p. 345). He referred to some studies to pinpoint that "knowledge which students had apparently acquired actually disappeared in a matter of months, probably indicating that the teaching had produced nothing more than pseudo-learning" (p. 346).

Since then, Truscott's argument has yielded several counter responses from a number of researchers (Bruton, 2009a, 2010; Ferris, 1999, 2004, 2006; Lyster, Lightbown, & Spada, 1999). Ferris (1999) was the first and argued that Truscott had limited and conflicting evidence and thus his conclusions were premature, so more research was needed before making a firm logical decision was made either to offer the CF or not. However, she agreed that evidence from well-designed studies was necessary before any firm conclusions could be drawn about the (in)effectiveness of error correction Ferris (1999, 2004). This call has resulted in a large number of studies exploring the effects of WCF on L2 learners' writing.

As a matter of fact, it seems that Truscott ignored a number of reasonable claims made by cognitive and socio-cultural theorists (Bitchener & Ferris, 2012). He did not consider the possibility that some learners who are ready, in Pienemann's terms (Pienemann, 1998) and socio-cultural terms (the ZPD), to acquire the form or structure targeted with WCF may be able to use the L2 form or structure with greater accuracy over time (Bitchener & Ferris, 2012). Bitchener and Ferris (2012) believed that this can occur irrespective of whether or not teachers understand, agrees, or disagrees with the processes underlying interlanguage development. Moreover, Truscott (1996) doubted two issues: (1) whether some teachers

are capable of providing consistent and correct CF and (2) whether some learners can understand what is meant by the CF. However, Bitchener and Ferris (2012) disagreed with Truscott and stated that the L2 teachers usually have little difficulty finding and correcting the learners' most problematic and most recurring linguistic errors as well as the errors which their learners are cognitively ready to respond to with WCF. Based on Bitchener and Ferris (2012), teachers can also consult commercially produced course books, designed for learners at different proficiency levels, because the linguistic forms and structures, in these books, have been selected on the basis of their appropriateness for learning at certain levels.

Although Truscott has continued to maintain his justifications for his position, he has suggested that written CF may be effective in certain limited situations (Truscott, 1999, 2004, 2007, 2009; Truscott & Hsu, 2008). Consequently, "the role of written CF in L2 development is an exciting and dynamic area of investigation, and, as such, is likely to continue engaging the energy and insights of established and emerging scholars" (Bitchener & Ferris, 2012, p. 27).

2.4. Research on WCF in L2 Writing

As Bitchener and Ferris (2012) explained, until the 1990s, almost all research on error correction in L2 writing focused on language instruction in ESL/EFL classes, not specifically on composition or writing classrooms. Beginning in the 1990s, because ESL writing and error treatment in L2 writing classes has turned into a distinct research field, researchers began to investigate ways in which writing teachers responded to L2 students' errors and the effects of these interventions. These studies examined short- and longer-term effects of error correction on student accuracy (Benson & DeKeyser, 2018; Bitchener, 2008; Bitchener et al., 2010, 2005; Bitchener & Knoch, 2008, 2009, 2010b, 2010a; Chandler, 2003; Ellis et al., 2008; Ferris, 1995a, 1997, 2006; Frear & Chiu, 2015; Karim & Nassaji, 2018; Kepner, 1991; Liu, 2016; Mawlawi Diab, 2015; Nassaji, 2011; Polio et al., 1998; Sadeghi et al., 2013; Santos et al., 2010; Sato, 2012; Semke, 1984; Y. Sheen, 2007, 2010b; Y. Sheen et al., 2009; Truscott, 2007; Van Beuningen et al., 2012), the impact of different correction techniques on immediate student revision (Ashwell, 2000; Ferris, 1997; Ferris & Roberts, 2001; Sheppard, 1992; Truscott & Hsu, 2008; Van Beuningen, De Jong, & Kuiken, 2008; Van Beuningen et al., 2012), descriptions and self-reports about how teachers respond to student error (Bailey & Garner, 2010; Ferris, 2006; Ferris, Brown,

Liu, & Stine, 2011; K. Hyland, 2013; Lee, 2008, 2009; Montgomery & Baker, 2007), and student views of teacher response, including error correction (Alkhatib, 2015; Chandler, 2003; Cohen & Cavalcanti, 1990; Cumming, 1995; Diab, 2005a, 2005b; Elwood & Bode, 2014; Ferris, 1995b, 2004; Ferris & Roberts, 2001; Grami, 2005; Hajian et al., 2014; Hamouda, 2011; Hedgcock & Lefkowitz, 1994, 1996; Horbacauskiene & Kasperaviciene, 2015; F. Hyland, 1998; Lee, 1997, 2004; Leki, 1991; Mahfoodh & Pandian, 2011; Montgomery & Baker, 2007; Westmacott, 2017; Zacharias, 2007).

2.5. Studies on the Effect of Unfocused WCF on L2 Accuracy Gains

As one of the pioneers, Semke (1984) compared the effects of four CF methods (writing comments and questions without corrections, marking all errors and correcting them, providing both positive comments and corrections, and indicating errors by "a code and requiring students to find corrections and then rewrite the assignment") (p. 195) with 141 German freshmen. The results indicated that student progress was improved by writing practice alone. The treatments did not increase the writing accuracy. Nonetheless, Guénette (2007), after analyzing the Semke's (1984) work, argued that the findings could have been affected by the presence of an incentive for some participants. The 'content' group was given score considering the number of words written, but the 'form' groups were graded on a ratio of errors to the number of written words. As a result, Guénette (2007, p. 50) suggested that 'content' group could have not been "worried about losing points, while the other three groups probably needed to write less for fear of making too many mistakes." Given these two issues (a real control group and an incentive), Semke's findings cannot be accepted as clear evidence for the absence of an effect for comprehensive WCF.

Then, Kepner (1991) measured the level of grammatical accuracy after assigning two types of WCF (error-corrections vs. message-related comments) to Spanish college students. "All sentence-level errors in the student writings in the surface-corrections group were identified, corrected, and explained via a brief rule or note." (p. 305) There was no significant difference in the number of errors between the two groups after a 12-week period; however, Ferris (2003, 2004) suggested that there was at least weak evidence of a beneficial effect for written CF because the correction group had fewer errors. Moreover, Ferris (2003) found a serious problem with the design of Kepner's study. Kepner had analyzed the sixth set of journal entries (out of eight) for comparison purposes, (and that) she did not look at the first set of journal entries to see where the students started out on

this type of writing. Therefore, as Ferris (2003) stated, there is no way of knowing whether the two groups of learners had the same initial level of accuracy. Further, as Ferris (2003) found, “there was no control for the length of the journal entries (which were written outside class), a variable that could affect both error and proposition counts” (p. 60). Moreover, Bitchener and Ferris (2012) argued that in Kepner’s study, the collected students’ journal entries as their writing were neither revised nor individually graded which received feedback consisted of that. Thus, it is unlikely that the learners in Kepner’s study were paying close attention to the received feedback and if the students had not paid attention to feedback, the findings would not be used as evidence that such feedback would be ineffective (Bitchener & Ferris, 2012). As a result, it is crucial to interpret the results of the study with caution because it is not clear evidence of the ineffectiveness of comprehensive WCF.

Some years later after Kepner, Truscott and Hsu (2008) investigated the effect of unfocused corrections on writing of 47 ESL learners and found that comprehensive CF made their learners able to improve the accuracy of a particular text during revision, but it did not lead to accuracy gains when writing a new text. Nevertheless, Bruton (2009) noted that the participants in Truscott and Hsu's (2008) study had very few errors during the pretest in the beginning of the study, so this could have affected the findings.

Van Beuningan et al., (2008) explored the effects of comprehensive CF on the writing of 66 L2 Dutch learners and found that learners who received comprehensive CF made significantly fewer errors in newly produced texts than students whose errors had not been corrected. However, Bitchener and Ferris (2012) found some shortcomings in Van Beuningan et al.'s (2008) study. Bitchener and Ferris (2012) stated that it is not clear whether Van Beuningan et al., (2008) controlled for additional linguistic input between giving feedback and the one-week delay in giving the immediate post-test, and whether the self-correction group was a real control group given that the learners’ attention was focused on accuracy when doing the self-correction and given that this occurred before the immediate post-test, one week later.

Recently, Bonilla López, Van Steendam, Speelman, and Buyse (2018) studied the potential of four comprehensive CF forms (DCF and metalinguistic codes, both for grammatical errors as well as DCF and metalinguistic codes, both for grammatical and non-grammatical errors) as editing and learning tools with 139 low-intermediate L2 writers, who consisted four experimental groups and a control group (self-correction). Results revealed that DCF

and codes were beneficial for improving learners' grammatical and non-grammatical accuracy during text revision; however, a long-term advantage (i.e., 4 weeks after feedback provision) was only found for direct corrections.

Then, Karim and Nassaji (2018) explored the short-term and delayed effects (after a two-week interval) of comprehensive WCF on L2 learners' revision accuracy and writing text. They investigated the efficacy of direct CF and two types of indirect CF, involving underlining only and the other underlining+metalinguistic cues. All the three feedback types significantly improved the revised texts. However, the accuracy improvements on new writing text, which was found for DCF and underlining+metalinguistic feedback types were non-significant.

2.6. Studies Comparing the Effect of Focused with Unfocused WCF on Accuracy Gains

Among the early studies which compared the effect of focused with unfocused WCF on accuracy gains, Ellis et al., (2008) can be mentioned. Ellis et al., (2008) compared the effectiveness of selective and comprehensive WCF on the accuracy of the English articles in written narratives of 49 Japanese university students. The researchers found that the CF was equally effective for the selective and comprehensive groups. However, Bitchener and Ferris (2012) believed that "the two types of feedback were not sufficiently distinguished from one another because article corrections were highly represented in both. Consequently, the findings need to be interpreted in this light" (p. 62).

On the other hand, Sheen et al., (2009) who also compared the effects of focused and unfocused WCF on the accuracy gains with 80 ESL intermediate students, found that the focused approach was more beneficial than provision of unfocused feedback when both were compared with the control group. Nonetheless, the researchers pointed out that the correction received by the unfocused CF group was rather unsystematic in nature; some errors were corrected, while others were ignored. Therefore, as Van Beuningen et al., (2012) stated, this unsystematic way of correcting, negatively affected the effect of unfocused CF in Sheen et al.,'s (2009) study.

In the recent decade, Frear (2010) compared the effectiveness of selective DCF for past tense forms, comprehensive DCF and a control group (no CF) on the accuracy of student's new piece of writing. The participants were 91 Chinese university students, who had been studying English for around 10 years. There was a range of proficiency in the classes, but

the vast majority could be classified as being at an intermediate level. Three conditions improved in accuracy; nonetheless, both the selective DCF and comprehensive DCF significantly improved the accuracy of the second writing task in comparison to the control group.

Then, Farrokhi and Sattarpour (2012) examined (1) whether direct WCF could improve the accuracy of the high-proficient L2 learners' writing performance with regard to the use of the English articles 'a' and 'the'; and (2) whether there were any differential effects in offering the focused DCF and the unfocused one on the accurate use of the English articles. 60 high-proficient participating EFL learners were assigned to three groups: focused WCF, unfocused WCF, and no feedback. It was found that both experimental groups outperformed control group in the post-test; additionally, the focused feedback proved to be more significantly beneficial than the unfocused one.

Next, Kassim and Ng (2014) compared the short and long term effects of selctive indirect and comprehensive indirect CF on the accurate use of prepositions by 90 upper-intermediate ESL learners' written work over 12 weeks. It was found both treatments improved the accuracy of the the posttests, yet no significant difference was revealed between the unfocused and focused CF groups.

Recently, Modirkhamene, Soleimani, and Sadeghi (2017) studied the effect of three feedback types (i.e., focused coded underlining, unfocused coded underlining, and no correction) on the written accuracy of 66 elementary EFL learners both in the short and long term. During 11 weeks, the target grammatical points "(past tense, countable/uncountable, and comparative adjectives) and all grammar errors" (p. 193) in the writing of the participants. Nevertheless, the 'no correction' group received qualitative comments. The focused feedback significantly improved the learners' accurate use of the target grammatical structures both in the post and delayed-post tests.

As the most recent publishd paper which compared the effect of focused with unfocused WCF on accuracy gains, Kristia (2018) can be mentioned. Kristia (2018) investigated the possible existence of the differences in the effects of focused and unfocused direct WCF on 12 Hong Kong ESL learners' writing. The focused feedback was provided only on indefinite and definite article errors whereas the unfocused feedback was given on all kinds of errors. In form of a pre-test–immediate post-test–delayed post-tests design, the groups wrote narrative stories and completed error correction tests. Both groups gained

improvements in both tests, showing the significant main effect of time. The mentioned two types of CF were equally effective.

2.7. Studies on the Effect of Unfocused WCF on L2 Accuracy and Complexity Gains

As the pioneers who studied the effect of unfocused WCF on both L2 accuracy and complexity gains, Robb et al., (1986) investigated the effect of four types of comprehensive CF (direct correction, the coded feedback, the uncoded feedback, and the marginal feedback) on both accuracy and complexity of narrative essays written by 134 Japanese college freshmen. Regarding the accuray, they found that the accuracy of all groups' writings improved and no difference was found between the effectiveness of CF types. The complexity of essays was analyzed based on "the ratio of additional clauses to total words written" and "additional clauses" (p. 90). The researchers reported that although the complexity of essays in all groups significantly improved, there was not any significant difference between them. As for an elaboration on the measures used by the researchers, it should be pinpointed that they used their own accuracy and complexity measures which included 1 subjective and 18 objective measures of writing ability. The following are the details of their complexity measures (Robb et al., 1986, p. 87):

- A holistic range of writing ability
- The Usage Correctness Score [which was first used in Brodkey and Young (1981)]
- The total of words written
- The total number of additional clauses embedded in T-units
- The total number of error-free T-units
- The total number of T-units written
- The ratio of words in error-free T-units to total error-free T-units
- The ratio of error-free T-units to total T-units
- The ratio of error-free T-units to total clauses
- The ratio of error-free T-units to total words written
- The ratio of words in error-free T-units to total word
- The number of words per T-units
- The ratio of words in error-free T-units to total T-unit
- The number of error-free clauses
- The ratio of total clauses to total words written
- The ratio of additional clauses to total words written

- The total number of clauses written
- The total number of words in error-free T-units
- The ratio of error-free clauses to total clauses

According to Van Beuningen et al., (2012), the absence of a no-correction control group can be the shortcoming of Robb et al's study.

Later, Sheppard (1992) examined the effects of two types of feedback to essays: "discrete item attention to form and holistic feedback on meaning" (p. 103) with 26 college freshmen from the Caribbean, Latin America, Europe, and Asia. "The ratio of subordinations to the total number of sentences" (p. 106), which is an indirect measure of complexity, was analyzed. Sheppard found that 'discrete item attention to form' had a negative effect on the structural complexity, but this finding was not significant. To analyze the accuracy of the essays, "1) The percentage of correct verb forms", except for "the correctness of negating devices" and "2) The percentage of correct sentence-boundary markers", (p. 106) except for the fragments, were considered. It was revealed that the use of holistic feedback on content/meaning improved the grammatical accuracy more than the other utilized feedback. Sheppard's (1992) study had some shortcomings. First, like Robb et al.'s (1986) study, the Sheppard's one also did not include a no-correction control group. Furthermore, Ferris (2003, 2004) pointed out that there was an absence of inter-rater reliability checks on the coding of the data, which makes us cautious about accepting the findings.

Then, Chandler (2003) compared the direct and indirect comprehensive CF with 31 music majors, from Korea, Japan, China, and Taiwan, who were required to write almost 25 pages about their lives as well as a book review. Chandler (2003) utilized holistic ratings and concluded that the WCF did not have any effect on the complexity of L2 learner writing. Chandler also found that "both direct correction and simple underlining of errors are significantly superior to describing the type of error, even with underlining, for reducing long-term error" (p. 267) However, as Van Beuningen et al., (2012) truly stated, one problem with the Chandler's study was that the utilized holistic ratings; the fact that holistic ratings did not show any changes is not necessarily a convincing proof "that the linguistic complexity of learners' writing did not change either" (p. 9).

More recently, Van Beuningen et al., (2012) compared the effect of direct and indirect comprehensive CF on 168 multilingual teenage students' accuracy and complexity of

writing. The participants were from different countries and thus were multilingual; they had limited L2 proficiency. As Van Beuningen et al., (2012, p. 18) explained, to measure structural accuracy and complexity, they used a subordination index: the number of subordinate clauses as a percentage of the total number of clauses (i.e., [number of subclauses/total number of clauses]×100). … Lexical diversity was calculated using Guiraud's [1954] Index, a type-token ratio that corrects for text length (types/√tokens).

Finally, it was revealed that both direct and indirect comprehensive CF improved the written accuracy. Nevertheless, as for the effect of feedback types on types of errors, it was found that only DCF led to grammatical accuracy improvements in new writing, whereas indirect CF was advantageous for the non-grammatical accuracy gains. As for the structural complexity, the researchers reported that both direct and indirect unfocused WCF did not lead to simplified writing.

Recently, Fazilatfar, Fallah, Hamavandi, and Rostamian (2014) investigated whether the comprehensive WCF was effective on the syntactic and lexical complexity of 30 Iranian EFL learners' writing at advanced level, who enrolled in conversational classes. Lu's (2010, 2012) L2 Syntactic and Lexical Complexity Analyzer were used to evaluate the syntactic and lexical complexity of the students' writing. To assess the syntactic complexity, the researchers calculated the length of production at mean length of sentence (MLS), and a dependent clause ratio (dependent clauses per clause, or DC/C). The study found a significant positive effect on both syntactic and lexical complexity on the group that had received the comprehensive WCF.

2.8. Studies on the Effect of Direct WCF

Among the pioneers who studied the effect of direct WCF, Bitchener et al., (2005) compared the effects of two types of feedback (direct, explicit WCF) either combined with or without student–researcher individual conferences on the accuracy of three types of error (prepositions, the past simple tense, and the definite article) over a 12 week period. The paticipants were 53 post-intermediate adult learners. It was found that the combination of written and conference feedback significantly improved the accurate use of the past simple tense and the definite article; however, after considering the three types of errors as a single group, it was revealed that the provided feedback types did not improve the overal accuracy of the written text. The researchers concluded "that L2 learners, in the process of

acquiring new linguistic forms, may utilize them with accuracy on one occasion but cannot do so on other similar occasions" (p. 191).

It should be noted that several studies have compared the effect of direct CF with indirect one (Bitchener & Knoch, 2010a; Chandler, 2003; Ferris, 2006; Frantzen, 1995; Lalande, 1982; Robb et al., 1986; Van Beuningan et al., 2008). Some researchers have indicated no differences between the two CF types (Frantzen, 1995; Robb et al., 1986), others have revealed that the indirect CF was more advantageous (Ferris, 2006; Lalande, 1982), and whereas there exist studies which have shown that DCF was more effective (Bitchener & Knoch, 2010a; Chandler, 2003; Van Beuningan et al., 2008).

2.9. A Study on the Effect of Metalinguistic CF on L2 Accuracy Gains

Azizi, Behjat, and Sorahi (2014) studied the effects of two types of metalinguistic feedback: error codes feedback and description feedback with 69 Iranian female high school students who were randomly assigned to three groups: no-feedback, error code feedback, and description feedback. Results showed that the metalinguistic description CF improved the writing of the participants.

2.10. Studies Comparing the Effect of Direct CF with Metalinguistic CF on Accuracy Gains

More than a decade ago, as one of the pionneers who compared the effect of DCF with metalinguistic CF on accuracy gains, Sheen (2007) first compared the effects of two types of WCF (direct-only & direct metalinguistic) on the acquisition of articles; then, she investigated "the extent to which language analytic ability mediates the effects of CF" (p. 255). The participants were 91 adult intermediate ESL learners. There was also a control (i.e., no-correction) group. It was revealed that both interventions significnalty improved the immediate posttests in comparison to the control group, yet the direct metalinguistic group outperformed the direct-only CF group in the long term. A significantly positive association was also found between students' improvement and their language analysis aptitude. In other words, the higher the students' language analytic ability, the more effective the WCF. Further, "language analytic ability was more strongly related to acquisition in the direct metalinguistic group than in the direct-only group" (p. 255).

A year later, Bitchener (2008) investigated the effects of WCF types on the accuracy of English articles "a" and "the" in the descriptive writing of 75 low-intermediate ESL

students. There were 4 groups: DCF, written and oral ME, DCF combined with written ME, no-CF. The participants wrote three pieces of writing: pre-test, immediate post-test, and delayed post-test. It was found that the WCF improved the accuracy in the immediate post-test, which was retained in the delayed-post test 2 months later. Then, Bitchener and Knoch (2008) extended the Bitchener's (2008) study and included an additional 69 low-intermediate learners (totally 144 international and migrant ESL students). It was revealed all three WCF types improved the accuracy of the posttest, which was retained in the delayed-posttest over seven weeks; in addition, no difference was found between the performance of migrant and international students. Bitchener and Ferris (2012) stated that because in Bitchener and Knoch's (2008) study, no difference was found between the same three treatment combinations, it is likely that the larger sample size removed the difference between the groups in Bitchener's (2008) study. Then, a year later Bitchener and Knoch (2009) published the result of their six-month study with 39 low-intermediate ESL learners. The researchers investigated the previously studied three WCF types on English articles "a" and "the", but this time the participants wrote four pieces of writing: "pre-test, immediate post-test, and two delayed post-tests" (p. 322). No difference on accuracy was revealed between the three WCF types, which suggests "that the provision of error correction alone may be sufficient for learners at a low-intermediate proficiency level" (p. 322). Bitchener and Knoch (2010b) published their other more longitudinal study over a period of 10-month on the effects of the above-mentioned WCF types on two English articles 'a' and 'the' with 52 low-intermediate ESL students, but this time the participants wrote five pieces of writing: "pre-test, immediate post-test, and three delayed post-tests" (p. 193). The treatment groups outperformed the control one on all post-tests and no difference was revealed between the three WCF types in their efficacy. Bitchener and Ferris (2012) believed that the special significance of this Bitchener and Knoch's (2010b) "finding was its investigation over a 10-month period and therefore its longitudinal measurement of the effectiveness of different types of CF on accuracy retention" (Bitchener & Ferris, 2012, p. 69).

Then, Shintani, Ellis, and Suzuki (2014) compared the effects of DCF and ME on 214 Japanese university students' accuracy of use of indefinite article and the hypothetical conditional. The participants were considered as pre-intermediate learners. Both types of feedback were provided with and without a revision opportunity in the form of rewriting. The feedback improved the accuracy of the hypothetical conditional but not of the

indefinite article. It was also proved that the lasting effect of the DCF was longer than one for the ME. Further, the revision opportunity positively affected the efficacy of the feedback. The targetd structures in this study, at which the WCF was directed, vary in saliency and complexity. Based on the findings, it can be suggested, in this situation, the "learners are likely to focus on the structure that contributes more to the global meaning of the text" (p. 103). It can also be concluded tht the DCF is more beneficial than the ME with regard to the complex syntactical structure.

Also, Gholaminia, Gholaminia, and Marzban (2014) compared the direct CF with meta-linguistic CF on 91 Iranian university students' paragraph writing. The researchers found that the meta-linguistic code-correction proved to be more beneficial than the DCF group.

Next, Stefanou and Révész (2015) via a classroom-based study with two treatment sessions and 89 Greek EFL learners, compared the efficacy of DCF as well as DCF+ME on the use of English definite article "in relation to learner differences in grammatical sensitivity and metalinguistic knowledge" (p. 263). There was also a control group who received no feedback. The study employed a pretest–posttest–delayed posttest design. To measure the learners' ability development, "a text summary and a truth value judgment test" (p. 263) were utilized. It was found that receiving the DCF was advantagous in comparison with no feedback; however, no clear benefit was found for ME. It was also proved that the greater the students' grammatical sensitivity and metalinguistic knowledge, the more likely the achieved gains in the DCF.

Then, Mawlawi Diab (2015) compared the effects of two types of WCF: DCF+ metalinguistic feedback and only metalinguistic feedback on the written accuracy of the ESL students. The feedback was directed at pronoun agreement errors and lexical errors. There was also a control group that self-edited their errors. All groups were required to revise their errors after receiving the feedback and before writing the next essay. Dirct metalinguisitic feedback significantly improved the pronoun agreement errors on the immediate post-test, whereas no significant difference was revealed in lexical errors. Nevertheless, at the delayed post-test, no significant difference was indicated among the groups in pronoun agreement errors, yet a significant difference was proved in lexical errors of the direct metalinguistic group.

2.11. Studies on the Potential Mediating Factors in WCF Efficacy

Participant classification, in WCF literature, is identified as one of the variables which might determine the effectiveness of the WCF, but the issue is not sufficiently explored (Ferris, 2003, 2004). For instance, considering the differences between the EFL and ESL contexts in the purposes for which the learners are writing and in the pedagogical contexts in which they have learned their L1 and L2 literacy, Hedgcock and Lefkowitz (1994) stated that EFL students may be less motivated than ESL ones to pay attention to WCF because they probably study to fulfil a kind of qualification requirement, not to improve their English in terms of its accuracy and clarity, so that they can become fully active members of an English-speaking community.

The other identified participant variable that may affect the efficacy of WCF is the learners' educational background (Hedgcock & Lefkowitz, 1994). According to Hedgcock and Lefkowitz, (1994), in product-oriented instruction, the students receive summative feedback to justify a score and the purpose is not to help the learners revise their texts and improve the accuracy of their new pieces of writing as opposed to formative feedback (as a result of process-oriented instruction where feedback is provided with the opposite intention); therefore, in product-oriented instruction, the learners may not be highly motivated to pay attention to WCF.

Another background factor which has been identified as a variable which might have an effect on the efficacy of WCF is the level of exposure the student has had to formal and acquired knowledge of the L2 (Ferris, 1999). For example, the EFL students may have a strong explicit knowledge of terms and rules of English grammar but less of an acquired awareness of how to apply such knowledge to their own writing. On the other hand, the students who have not learned the L2 via formal instruction but who have picked it up through participation in informal, conversational contexts may depend on their acquired L2 knowledge "to self-correct their writing, but have limited or perhaps no awareness of how to reference the formal system of grammar elicited by WCF" (Bitchener & Ferris, 2012, p. 71).

To investigate these issue, Bitchener and Knoch (2008) divided their 144 participants into two groups: 75 international (visa) students and 69 migrant ones. The international students had studied the L2 in an EFL context and were used to receiving WCF and focusing on accuracy in their L2 classrooms, while the migrant students were "less accustomed to formal classroom learning and therefore potentially less attuned to focusing

on WCF" (p. 410). Errors in the use of the same English articles, targeted in other studies referred to above, received WCF. There was also a control group that did not receive WCF. In spite of the earlier suggestions (e.g., Ferris, 1999; Hedgcock & Lefkowitz, 1994), Bitchener and Knoch (2008) found that there was no difference in the extent to WCF contributed to the improvement of the migrant and international students' writing. Based on Bitchener and Ferris (2012), the result might be because neither of the two groups consisted of learners "exclusively from one of these backgrounds". For example, "migrant students may or may not have had formal instruction" in the L2. "International visa students may or may not have had opportunities to study the L2 in an English-speaking environment (e.g., during term holidays)" (Bitchener & Ferris, 2012, p. 71).

The other factor which might have an influence on the degree to which the learners are able to benefit from the WCF is the learners' educational level because it is presumed that learners' educational level can show their level of metalinguistic awareness, and it has been suggested that students with lower levels of (meta)linguistic competence might be less able to correct their own errors based on WCF (Ferris, 2004; K. Hyland & Hyland, 2006; Y. Sheen, 2007). Nonetheless, Van Beuningen et al., (2012) investigated the influence of students' educational level on CF efficacy, assuming the higher level students might be better equipped with metalinguistic knowledge than their lower level peers. The findings did not confirm these expectations; even though the higher level students did outperform the lower level learners on the different linguistic measures used in the study, no significant interaction was found between the effectiveness of the CF treatments and learners' educational level. Nonetheless, as Van Beuningen et al., (2012) stated, the difference between the levels included in this study was not large enough, which could be the reason why they did not find a significant interaction between students' educational level and WCF efficacy. In addition, the participants in their study started learning Dutch at an early age, and the L2 was the means rather than the goal of instruction; therefore, the "learners' level of metalinguistic awareness might have been too low for them to really fully benefit from indirect CF, even for the higher level pupils" (p. 34).

In addition to the mentioned factors, a task's topic, also proved to contribute to a learner's written performance (e.g. Schoonen, 2005). Specifically, the level of knowledge about, interest in, and familiarity with a certain topic have proven to affect writing performance (e.g. Benton, Sharp, Corkill, Downey, & Khramtsova, 1995; McCutchen, 1986). Another factor that might add to this task-effect is the conceptual complexity of a writing task

(Kuiken & Vedder, 2008). Such factors might also affect the extent to which a learner benefits from the received WCF on the performance on a given task or topic. If this would be the case, it would compromise the generalizability of findings found in studies on CF (Van Beuningen, 2011).

2.12. The Necessity to Do Further Studies

As it was already mentioned, the majority of feedback related studies using measures of accuracy explored the effects of focused or selective feedback on EFL writing (Bitchener, 2008; Bitchener et al., 2010, 2005; Bitchener & Knoch, 2008, 2009, 2010b; Y. Sheen, 2007, 2010a; Y. Sheen et al., 2009), whereas "evidence on the language learning potential of unfocused CF, which involves comprehensive correction of every error in students' writing, is scarce" (Van Beuningen et al., 2012, p. 5).

Furthermore, although many studies focused on the effects of WCF on accuracy, to date, only very few studies have investigated the effect of WCF strategies on developing syntactic complexity in L2 writing (Chandler, 2003; Robb et al., 1986; Sheppard, 1992; Van Beuningen et al., 2012). These few studies on the effect of WCF on written complexity also have revealed conflicting findings (Bitchener & Ferris, 2012). Moreover, with regard to the effect of WCF on syntactic complexity, as Guénette (2007) argued, "the success or failure of corrective feedback will depend on the classroom context, the type of errors students make, their proficiency level, the type of writing they are asked to do, and a collection of other variables" (pp. 51-52). In addition, the results of the studies on the effect of unfocused WCF on L2 accuracy and complexity gains are not comparable because of the differences they have with regard to their treatment period, methodology, measurement instruments, genre of writing task, as well as their participants' conditions (e.g., age, proficiency level, L1 background, L2 learning goals, etc.). For example, as Ortega (2015) stated, L1 is "a moderating variable of L2 syntactic complexity" (p. 84). The moderating effect of L1 on L2 syntactic complexity was also proven in Lu and Ai's (2015) study. Further, genre/task/content also play roles in syntactic complexity in writing and L2 proficiency is "a powerful source of influence that modulates syntactic complexity" (Ortega, 2015, p. 88). Therefore, the effect of WCF strategies on L2 syntactic complexity is a matter which continues to be a subject of controversy and requires more meticulous studies.

CHAPTER III

METHODOLOGY

3.1. Introduction

This study aimed at investigating the effects of two different types of WCF strategies: DCF and ME on L2 learners' not only syntactic accuracy but also syntactic complexity in writing both in the short and long term. In this chapter, detailed information about the participants, procedures, instrumentation, research design, and statistical analyses used in the study is presented.

3.2. Research Questions

The study addressed the following research questions:

1) Is there any significant difference between the effects of DCF and NF (i.e., no feedback) on developing learners' syntactic accuracy in writing in the short and long term?
2) Is there any significant difference between the effects of ME and NF on developing learners' syntactic accuracy in writing in the short and long term?
3) Is there any significant difference between the effects of DCF and ME on developing learners' syntactic accuracy in writing the short and long term?
4) To what extent are the comprehensive DCF and ME effective on developing EFL learners' written syntactic complexity, relative to each other and to NF (i.e., no feedback) in the short and long term?

3.3. Research Setting

The study was carried out at Antalya International University in Antalya, Turkey. The criterion for choosing this university was only based on the fact that the researcher had the

opportunity of teaching Essay Writing in preparatory university courses in this university. Therefore, the sampling method used in this study was convenience sampling.

3.4. Research Participants

The participants were 90 Turkish EFL learners between the ages of 18 to 23, who were selected among 127 students based on the results of their quick placement test (QPT). They were assigned to three groups; therefore, there were 30 participants in each group. The decision for 30 participants per group was made based on Hatch and Lazaraton's (1991) suggestion that a minimum of 30 participants for quantitative data analyses is required as the size may produce a distribution of a range of scores. The participants took Essay Writing courses as part of preparatory university courses at Antalya International University in Antalya, Turkey. There were 19 females and 11 males in the DCF group. The ME group included 18 females and 12 males. The NF group included 13 females and 17 males; thus, among 90 participants, there were 50 females and 40 males. They all had passed the elementary and intermediate writing courses, and enrolled in upper-intermediate writing prep-university courses.

3.5. Research Design

Bitchener (2008) stated that "if a post-test is to be a valid measurement of progress, a comparable pre-test needs to be included in the research design" (p. 108). Moreover, by including a delayed post-test in the design of the study, the validity of the findings (i.e., the effect of the feedback) will be sustained (Shintani & Ellis, 2013). Therefore, the present study used a pretest-treatment/control-posttest-delayed posttest design to determine the gains in both syntactic accuracy and syntactic complexity. Additionally, as it was mentioned in the Delimitations, the non-random convenience sampling method was used in this study, so it is a quasi-experimental one. Following Storch's (2010) recommendation, the study was done in real classrooms so that the feedback was provided in "the context of an instructional program, with ecologically valid writing tasks" (p. 43). In addition, following Guénette's (2007) recommendation, the groups in this study had the same teacher; the activities, courses, and writing topics were similar as well. This was done to control the possible effects of every other design parameter, except the feedback types.

Furthermore, in both experimental groups as well as the control group, to hold the learners responsible for their learning (Guénette, 2012), they had to revise their essays based on the

teacher's feedback. In addition, revision of the written texts after receiving feedback is perhaps a beneficial and essential step toward the long-term acquisition of a specific feature (Ferris, 2004, 2010; Guénette, 2012; Sachs & Polio, 2007).

As the final point, it should be noted that because 127 students had enrolled in upper-intermediate writing courses, it was unlikely to exclude them from the classes, so the researcher decided to work with all the learners, but for the purpose of the research, the scores (pretest, posttest and delayed-posttest) of the learners who were homogeneous were considered. In other words, the students who were not homogeneous were discarded from the research although they were present in classes.

3.5.1. Variables in the Study.

There were four main variables in the study: two main dependent variables and two independent variables.

The dependent variables were the gain in syntactic accuracy and syntactic complexity measured in the participants' production from the pretests to the posttests and delayed posttests. Syntactic complexity, as one of the dependent variables, included ten syntactic complexity measures: Mean length of clause (MLC), mean length of sentence (MLS), mean length of T-unit (MLT), clauses per sentence (C/S), clauses per T-unit (C/T), complex T-units per T-unit (CT/T), dependent clauses per clause (DC/C), coordinate phrases per T-unit (CP/T), complex nominals per T-unit (CN/T), and verb phrases per T-unit (VP/T). The independent variable was the comprehensive WCF type, (DCF and ME, in comparison to no feedback on syntactic errors). Additionally, the participants' English proficiency level and writing ability (i.e., written syntactic accuracy and complexity) were also considered as the control variables.

3.6. Instruments

To conduct the study, the following instruments were utilized: Student biodata information questionnaire, Quick Placement Test (QPT), class writing tasks, pretest, posttest, and delayed posttest. Additionally, in order to measure the syntactic accuracy, the formula which was also employed by Chandler (2003) as well as Truscott and Hsu (2008) was utilized. Additionally, to calculate the syntactic complexity, Lu's (2010) web-based L2 Syntactic Complexity Analyzer, entitled, 'Web-based L2SCA: Single Mode' was employed.

3.6.1. Students' Biodata Information.

The students' biodata information was obtained by an open-ended questionnaire, which asked the participants about their age, gender, English learning experiences (experiences of visiting English-speaking countries) and the length of time they had been studying or dealing with English (See Appendix 1 for more information about the students).

3.6.2. Quick Placement Test.

QPT has been developed by Oxford University Press and Cambridge ESOL for placement testing and examination screening (Geranpayeh, 2003). The QPT is divided into two parts and contains 60 questions. The complete test is attached to Appendix 2.

3.6.2.1. Quick Placement Test Reliability and Validity.

As Geranpayeh (2003) stated, the QPT, developed by Oxford University Press and Cambridge ESOL (English for Speakers of Other Languages), is a reliable test of finding a student's level of English language proficiency. It is reliable because all of the items in QPT have been through Cambridge ESOL quality control procedures. The QPT reports test results as a band on the ALTE (Association of Language Testers in Europe) 5-level scale and Council of Europe levels, which makes the results potentially much more useful.

Furthermore, to the date of Geranpayeh's (2003) study, the test had been validated in 20 countries by more than 6000 students. There were also some validation phases in the course of QPT development, which are explained in Geranpayeh's study on pages 2 and 3. The following are the mentioned phases in brief:

Phase 1: Students from different countries took the electronic version of the test and one of the two paper-and-pen tests. They also completed a questionnaire indicating the extent to which they were comfortable using a computer. Teachers also provided detailed feedback on the look and feel of the paper-and-pen tests and some of the items in the electronic version of the QPT, and how accurate the tests were in terms of identifying the current level of their students. As a result of this, the required modifications were made on paper-and-pen tests as well as the QPT database.

Phase 2: With the format of the tests confirmed, the second phase of activity concentrated on determining score equivalence between the electronic version of QPT and the paper-and-pen versions, and also between two successive administrations of the QPT. The aim

was to assess how consistently students were located on the same score and what degree of error was associated with these scores. As a result of these actions, the SEM of the test is around 4 and the typical reliabilities reported during the trial phase were close to .9 for the 60-item test and .85 for the 40-item test.

In spite of the mentioned point, as the findings were related to year 2003, the QPT was piloted before being used in the present study and the reliability of the test was calculated on Cronbach alpha scale.

3.6.3. A Pretest, a Posttest, a Delayed-Posttest, and Classroom Writing Tasks.

Adopting topics of the Writing Task 2 from the General Module of IELTS in this study was because of the following reasons:

To establish the *content validity, face validity,* and *test comparability,* first, the learners' background knowledge of the writing topics was really important to consider. The participants in this study were students of different majors. The researcher managed to find writing topics with which the participants were familiar.

In addition, as Mackey and Gass (2005) argued, researchers who do research into writing need to select appropriate topics based on the same form of essays because the pretest must be comparable in difficulty to the posttest and delayed-posttest. As a result, the exam topics in all pretest, posttest, delayed-posttest, and the writing tasks of this study were all argumentative type (See Appendix 3). The learners in the present study had already learned how to write argumentative essays, so they knew about the format of the test.

Finally, as for the *criterion-related validity* of the test, the samples from the writing tasks of the General Module of IELTS were used, so that the utilized tests and tasks in this study can be comparable to a standardized writing test.

3.6.4. Syntactic Accuracy Measure.

In order to measure the syntactic accuracy, the following formula was utilized; it was also employed by several previous researchers (Chandler, 2003; Truscott & Hsu, 2008): [total number of syntactic errors/total number of words] × 100.

3.6.5. Syntactic Complexity Measure.

To calculate the syntactic complexity, Lu's (2010) web-based L2 Syntactic Complexity Analyzer, entitled, 'Web-based L2SCA: Single Mode' was utilized. This analyzer can be downloaded from Lu's webpage at http://aihaiyang.com/software/l2sca/single/. To assess the syntactic complexity, the automatic approach was used "because it affords speed, flexibility, and reliability" (Crossley & McNamara, 2014, p. 69). Moreover, human raters are likely to be subjective and also they need training, time to score, as well as monitoring, all of which utilize resources (Higgins, Xi, Zechner, & Williamson, 2011). This analyzer "produces frequency counts of nine linguistic units in the text—word, sentence, clause, dependent clause, T-unit, complex T-unit, coordinate phrase, complex nominal, and verb phrase—and generates 14 indices of syntactic complexity for the text" (Yang, Lu, & Weigle, 2015, p. 58). In the present study, 10 out of 14 indices were analyzed. To see these 10 measures, see Table 3.1.

Table 3.1

Ten Syntactic Complexity Measures

Measure	Code	Definition
Type 1: Length of Production		
1. mean length of clause	MLC	number of words divided by number of clauses
2. mean length of sentence	MLS	number of words divided by number of sentences
3. mean length of T-unit	MLT	number of words divided by number of T- units
Type 2: Sentence Complexity		
4. clauses per sentence	C/S	number of clauses divided by number of sentences
Type 3: Subordination		
5. clauses per T-unit	C/T	number of clauses divided by number of T-units
6. complex T-units per T-unit	CT/T	number of complex T-units divided by number of T-units
7. dependent clauses per clause	DC/C	number of dependent clauses divided by number of clauses
Type 4: Coordination		
8. coordinate phrases per T-unit	CP/T	number of coordinate phrases divided by number of T-units
Type 5: Particular Structures		
9. complex nominals per T-unit	CN/T	number of complex nominals divided by number of T-units
10. verb phrases per T-unit	VP/T	number of verb phrases divided by number of T-units

It should be noted that because Ortega (2015) and Yoon (2017) stated that genre/task/content play roles in syntactic complexity in writing, to control their probable mediating effects, the participants were required to write essays selected from IELTS questions of argumentative type for the pre- and posttests, delayed posttest as well as the class writing tasks.

3.7. Data Collection Procedure

At the beginning of the study, all the participants were required to fill out a questionnaire concerning their personal information. It needs to be mentioned that the participants were ensured that their anonymity would be protected.

During the first week, to control the L2 proficiency level of the participants and to ensure that they were homogeneous, first, they were given the pen-and-paper version of QPT. To decide on the students' language level, Geranpayeh's (2003) guideline was used. Those who got scores ranging from 40 to 47 out of 60 (i.e., the upper-intermediate level), based on Geranpayeh (2003), were selected for the IELTS writing test, which was used to assess the syntactic accuracy and syntactic complexity in the written texts; this information was used both to ensure the homogeneity of the students and as the pretest. The students who were homogeneous in their writing ability were assigned to three groups, namely DCF (n = 30), ME (n = 30), who received feedback on their grammatical errors and a control group, namely, 'No Feedback', abbreviated as NF (n = 30) who was provided with feedback only on content, orthography, and organization, but not on grammatical errors because as Ferris (2004) stated, it seems almost unethical to single out a group for no feedback.

It should be noted that because 127 students had enrolled in upper-intermediate writing prep-university courses, it was unlikely to exclude them from the classes, so the researcher decided to work with all the learners, but for the purpose of the research, the scores (pretest, posttest and delayed-posttest) of the students who had the necessary criterion (i.e., they were homogeneous) were considered. In other words, the students who were not homogeneous were discarded from the research although they were present in classes.

The treatment/control period lasted for 4 more weeks. Every week the participants in each group wrote an argumentative essay in class and turned it in; the teacher-researcher provided the specified feedback to each group based on Ellis' (2009) examples. Then the students were required to revise their corrected text. Next, on the 1st session of Week 6, the posttest was administered. No work on writing was done during the 6th and 7th weeks.

Then, in the 8th week, the delayed posttest was implemented. Two points should be noted: (1) No student took the pre- and posttests and delayed posttests twice; and 2) The topics of essays were different in the mentioned tests. However, the topics were the same for all three groups.

3.8. Data Analysis

The main sources of data for this study were the QPT as well as the essays written by the participants during the pretests, posttests, and delayed-posttests. The accuracy of the essays was analyzed based on the following formula: [total number of syntactic errors/total number of words] × 100, while to calculate the syntactic complexity, Lu's (2010) web-based L2 Syntactic Complexity Analyzer, entitled, 'Web-based L2SCA: Single Mode' was utilized.

The Cronbach alpha indices were calculated as inter-rater reliability coefficients.

1. To find out whether the groups were homogenous, a one-way ANOVA test was run to compare the three groups on both QPT and pretest of accuracy. Furthermore, in order to find out whether the three groups were homogenous with regard to the syntactic complexity of their essays, a one-way between-groups multivariate analysis of variance (MANOVA) was performed on nine indices (MLC, MLS, MLT, C/S, C/T, CT/T, DC/C, CN/T, and VP/T), but because the correlation coefficient between the CP/T and other indices were mostly below .10, which is considered as very small correlation and this index could make the data violate the assumption of linearity, this index (i.e., CP/T) was removed from MANOVA and then a separate ANOVA test was performed on it.
2. Independent t-tests were run to compare the DCF and NF groups on the posttest and delayed-posttest of syntactic accuracy in writing.
3. Independent t-tests were run to compare the ME and NF groups on the posttest and delayed-posttest of syntactic accuracy in writing.
4. Independent t-tests were run to compare the DCF and ME groups on the posttest and delayed-posttest of syntactic accuracy in writing.
5. In order to explore the short and long-term effect of the WCF types on syntactic complexity indices (MLC, MLS, MLT, C/T, CT/T, DC/C, CN/T, and VP/T), two one-way between-groups MANOVAs (one for the posttest and one for the delayed-posttest) were performed, but because the correlation coefficient values found for the C/S and CP/T were mostly small, and these indices could make the data violate the assumption of linearity,

these indices (i.e., C/S and CP/T) were removed from MANOVA and then separate ANOVAs were performed on them.

CHAPTER IV

RESULTS AND DISCUSSIONS

4.1. Introduction

In order to find answers to the research questions, a series of statistical analyses were conducted. This chapter presents the results of data analyses conducted in the study as well as the discussions.

4.2. Inter-Rater Reliability Indices

The Cronbach alpha indices were calculated as inter-rater reliability coefficients. The indices ranged from a high of .97 for the delayed posttest of syntactic accuracy in the NF to .92 for the posttest of the DCF group.

Table 4.1

Inter-Rater Reliability Indices for Syntactic Accuracy

Groups	Tests on Syntactic Accuracy	Indices
DCF	Pretest	.94
	Posttest	.92
	Delayed-Posttest	.93
ME	Pretest	.96
	Posttest	.94
	Delayed-Posttest	.94
NF	Pretest	.96
	Posttest	.96
	Delayed-Posttest	.97

4.3. The Normality Tests

Examining graphics to visually look for the normality, can help, but graphic methods are not conclusive (Larson-Hall, 2010). In this study, the assumption of normality was first examined through the graphic of histogram, but because histograms “can only be a general,

not exact, measure of the normality of a distribution", some numerical ways were also utilized as recommended by Larson-Hall (2010, p. 77). Considering the size of the groups in this study, the histograms can be considered normally distributed. See histograms for QPT in Appendix 4, histograms for syntactic accuracy measures (pretest, posttest, and delayed-posttest) in Appendix 5, and histograms for pretest, posttest, and delayed-posttest of syntactic complexity in Appendices 6, 7, and 8, respectively.

4.3.1. Numerical Methods of Assessing Normality for QPT and Syntactic Accuracy.

In the present study, skewness statistics, kurtosis statistics, the ratio of skewedness and kurtosis over their respective standard errors were utilized as the numerical methods of assessing the normality (Field, 2013; Larson-Hall, 2010; Pallant, 2013; Phakiti, 2010). Phakiti (2010) stated that in studies of applied linguistics, "values of skewness and kurtosis statistics within +/-1 suggest that the data set is normally distributed" (p. 45). Besides, based on Field (2013), if the outcomes of the ratio of skewedness and kurtosis over their respective standard errors are within the ranges of +/-1.96, the data enjoy normal distribution. All the tests enjoyed normal distribution. That is why they were analyzed through the parametric tests. The table containing the detailed information of normality tests for QPT and syntactic accuracy is included in Appendix 9.

4.3.2. Numerical Methods of Assessing Normality for Pretest of Syntactic Complexity.

Based on Field (2013), the outcomes of the ratio of skewedness and kurtosis over their respective standard errors were within the ranges of +/-1.96; therefore, the data enjoy normal distribution. Further, Shapiro-Wilk was also used as another numerical method to test the normality status (Larson-Hall, 2010; Pallant, 2013) of the indices. Thode (2002) recommended the Shapiro-Wilk test as the best choice for testing the normality of data. All the *p*-values were greater than .05; therefore, it was concluded that the data enjoyed normal distribution. The tables containing the information of normality tests for pretest of syntactic complexity are included in Appendix 10.

The numerical tests for the posttest and delayed-posttest of syntactic complexity also revealed that the data enjoyed normal distribution. The outcomes of the ratio of skewedness and kurtosis over their respective standard errors were within the ranges of +/-

1.96 (Field, 2013). Besides, the values of skewness and kurtosis statistics were all within +/-1 (Phakiti, 2010). Moreover, all the *p*-values of Shapiro-Wilk tests were greater than .05, so it was concluded that the data were normally distributed (Larson-Hall, 2010; Pallant, 2013; Thode, 2002). All the relevant tables are included in Appendix 11.

4.4. Ensuring the Homogeneity of the Groups

4.4.1. Grammatical Knowledge.

A one-way between-groups analysis of variance (ANOVA) was conducted to explore whether the three groups (i.e., DCF, ME, NF) were homogeneous with regard to their grammatical knowledge, as measured by the QPT. Table 4.2 gives information about each group (number in each group, means, standard deviation, minimum, maximum, etc.).

Table 4.2

Descriptives for QPT

Groups	N	Mean	SD	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
DCF	30	44.03	1.65	.30	43.42	44.65	41	47
ME	30	44.00	1.50	.27	43.44	44.56	41	47
NF	30	44.10	1.62	.29	43.49	44.71	41	47
Total	90	44.04	1.57	.16	43.71	44.38	41	47

Table 4.3 indicates that the significance value (Sig.) for Levene's Test is greater than .05 (Sig. = .87), so the assumption of homogeneity of variance was met.

Table 4.3

Levene's Test for Homogeneity of Variances for QPT

Levene Statistic	df1	df2	Sig
.14	2	87	.87

As Table 4.4 reveals, there was not a significant difference among the three groups: F (2, .87) = .03, p = .97. In conclusion, the participants in three groups were homogeneous regarding their grammatical knowledge.

Table 4.4

ANOVA for the QPT

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.15	2	.07	.03	.97
Within Groups	221.66	87	2.54		
Total	221.82	89			

4.4.2. Syntactic Accuracy.

Then, another one-way between-groups analysis of variance (ANOVA) was conducted to explore whether the three groups (i.e., DCF, ME, NF) were homogeneous with regard to their syntactic accuracy, as measured by the essay writing test (i.e. pretest). Table 4.5 gives information about each group (number in each group, means, standard deviation, minimum, maximum, etc.).

Table 4.5

Descriptives for Pretest of Syntactic Accuracy

					95% Confidence Interval for Mean			
Groups	N	Mean	SD	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
DCF	30	12.19	.24	.04	12.10	12.28	11.66	12.67
ME	30	12.12	.26	.04	12.02	12.22	11.66	12.52
NF	30	12.08	.32	.05	11.96	12.20	11.48	12.70
Total	90	12.13	.28	.03	12.07	12.19	11.48	12.70

Table 4.6 indicates that the significance value (Sig.) for Levene's Test is greater than .05 (Sig. = .38); therefore, the assumption of homogeneity of variance was met.

Table 4.6

Levene's Test for Homogeneity of Variances for Pretest of Syntactic Accuracy

Levene Statistic	df1	df2	Sig
.97	2	87	.38

As Table 4.7 reveals, there was not a significant difference between the three groups: F (2, 87) = 1.22, p = .29. In conclusion, the participants in three groups were homogeneous regarding their syntactic accuracy in their pretests.

Table 4.7

ANOVA for Pretest of Syntactic Accuracy

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.19	2	.09	1.22	.29
Within Groups	6.85	87	.07		
Total	7.04	89			

4.4.3. Syntactic Complexity.

In order to find out whether the three groups were homogenous with regard to the syntactic complexity of their essays, a one-way between-groups multivariate analysis of variance (MANOVA) needed to be performed because there were more than one dependent variable (Larson-Hall, 2010; Pallant, 2013; Tabachnick & Fidell, 2013). This study included 10 dependent variables. MANOVA, however, has a number of assumptions that must be met. These include equal sample sizes, univariate normality, multivariate normality (outliers), multicollinearity and singularity, linearity, and finally, homogeneity of variance matrices (Pallant, 2013; Tabachnick & Fidell, 2013). As a result, these assumptions were first examined to be sure that they were not violated.

Fortunately, the study includes equal sample sizes in each group. Besides, the assumption of univariate normality has been already tested and found to be met.

To test for multivariate normality, Mahalanobis distance was calculated (Pallant, 2013; Tabachnick & Fidell, 2013). “Mahalanobis distance is the distance of a case from the centroid of the remaining cases where the centroid is the point created at the intersection of the means of all the variables” (Tabachnick & Fidell, 2013, p. 74). The maximum value for Mahalanobis distance for the DCF, ME, and NF were 26.12, 22.31, and 21.50, respectively. Pallant (2013) stated when there are 10 dependent variables, the critical value equals to 29.59. The maximum values for Mahalanobis distance for each group were less than the critical value, so it was safely assumed that there were no multivariate outliers.

Next, to check for multicollinearity and check the strength of the correlations among the dependent variables, a correlation test was run (Pallant, 2013; Tabachnick & Fidell, 2013). There were no correlations up around .8 or .9 (Pallant, 2013; Tabachnick & Fidell, 2013), so there was no reason for concern regarding multicollinearity. The table revealing the

results of correlation coefficients for the pretest of syntactic complexity is included in Appendix 12.

Nevertheless, the correlation coefficient between the CP/T and other indices were mostly below .10, which is considered as very small correlation; therefore, this index could make the data violate the assumption of linearity. Thus, following Pallant's (2013) recommendation, this index (i.e., CP/T) was removed from MANOVA and then a separate ANOVA test was performed on it.

Then, to assess the assumption of linearity, a matrix of scatterplots were generated between each pair of variables, separately for our groups (DCF, ME, and NF). The scatterplots (See Appendix 13) were roughly oval-shaped (Tabachnick & Fidell, 2013) or cigar-shaped (Pallant, 2013), so it was concluded that the variables were linearly related and the assumption of linearity was met.

Eventually, preliminary assumption testing to check for univariate and multivariate normality, multicollinearity, as well as linearity revealed no serious violations; thus, a one-way between groups multivariate analysis of variance (MANOVA) was performed to investigate the differences in syntactic complexity in three groups of DCF, ME, and NF. Nine dependent variables were used: MLS, MLT, MLC, C/S, VP/T, C/T, DC/C, CT/T, and CN/T. The independent variable was the type of WCF.

Based on Table 4.8, the assumption of homogeneity of variance-covariance was not met because the Sig. value was lower than .001. (.000 < .001), but there was a moderate violation, so as Tabachnick and Fidell (2013) recommended, a more stringent alpha level was adopted. For moderate violations, .025 has been recommended as the alpha level.

Table 4.8

Box's Test of Equality of Covariance Matrices for Pretest of Syntactic Complexity [a]

Box's M	190.183
F	1.800
df1	90
df2	20735.614
Sig.	.000

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

[a]. Design: Intercept + T

According to Table 4.9, all the Sig. values were over .05, indicating that the assumption of equality of variance for the variables was met.

Table 4.9

Levene's Test of Equality of Error Variances for Pretest of Syntactic Complexity [a]

	F	df1	df2	Sig.
MLS	.227	2	87	.797
MLT	1.126	2	87	.329
MLC	1.725	2	87	.184
C/S	.272	2	87	.762
VP/T	1.450	2	87	.240
C/T	1.276	2	87	.284
DC/C	1.206	2	87	.304
CT/T	.272	2	87	.762
CN/T	1.044	2	87	.356

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

[a]. Design: Intercept + T

As the second row of Table 4.10, which is for the treatment (i.e., types of WCF) indicates, the *p*-value is larger than the stringent alpha level (.49 > .025), so there was not a significant difference among the three groups on the combined dependent variables, *F* (18, 158) = .97, *p* = .49; Wilk's Lambda = .81; partial eta squared = .10. Given the mentioned information, the three groups were homogenous with regard to MLS, MLT, MLC, C/S, VP/T, C/T, DC/C, CT/T, and CN/T of the written essays.

Table 4.10

Multivariate Tests for Pretest of Syntactic Complexity [a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.997	3287.69[b]	9.000	79.00	.000	.997
	Wilks' Lambda	.003	3287.69[b]	9.000	79.00	.000	.997
	Hotelling's Trace	374.54	3287.69[b]	9.000	79.00	.000	.997
	Roy's Largest Root	374.54	3287.69[b]	9.000	79.00	.000	.997
T	Pillai's Trace	.194	.956	18.000	160.0	.51	.097
	Wilks' Lambda	.810	.972[b]	18.000	158.0	.49	.100
	Hotelling's Trace	.228	.988	18.000	156.0	.47	.102

	Roy's Largest Root	.199	1.770[c]	9.000	80.00	.08	.166

[a]. Design: Intercept + T
[b]. Exact statistic
[c]. The statistic is an upper bound on F that yields a lower bound on the significance level.

As it was already mentioned, the dependent variable of CP/T had been removed from the data before performing the MANOVA because it had made the data violate the assumption of linearity, and it was decided to run a separate ANOVA on it.

Table 4.11 indicates that the significance value (Sig.) for Levene's Test was greater than .05 (Sig. = .69), so the assumption of homogeneity of variance was met.

Table 4.11

Levene's Test for Homogeneity of Variances for Pretest of CP/T

Levene Statistic	df1	df2	Sig
.36	2	87	.69

As Table 4.12 reveals, there was not a significant difference between the three groups: F (2, 87) = .056, p = .94. In conclusion, the participants in three groups were homogeneous regarding the CP/T in their pretests.

Table 4.12

ANOVA for Pretest of CP/T

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.003	2	.001	.056	.946
Within Groups	2.226	87	.026		
Total	2.229	89			

4.5. Answer to the First Research Question

The first research question explored whether there is any significant difference between the effects of DCF and NF (i.e., no feedback) on developing learners' syntactic accuracy in writing in the short and long term.

4.5.1. The Short-Term Effect: Post-Test Result.

An independent *t*-test was run to compare the mean scores of the DCF and NF groups on the posttest of syntactic accuracy in writing in order to probe the effect of the DCF on the improvement of the syntactic accuracy of the students' essays shortly after the administration of the treatment. It should be noted that as the formula [total number of

syntactic errors/total number of words] × 100 was utilized for scoring the syntactic accuracy of the essays, the fewer errors the essays included, the smaller value (mathematical quantity) they were given, so the lower values reveal the existence of fewer errors.

It should also be mentioned that the assumption of homogeneity of variances was met (Levene's F = 2.66, p = .10 > .05). That is why the first row of Table 4.14, "Equal variances assumed" is reported. Based on the results displayed in Tables 4.13 and 4.14, it could be concluded that on average, the DCF group (M = 11.05) outperformed the NF group (M = 12.07). This difference was significant (t (58) = -14.81, p < .05); the p-value for this t was .000 (Sig (2-tailed) = .000 < .05); however, the magnitude of the differences in the means (mean difference = -1.01, 95% CI: -1.15 to -.88 was below medium based on Cohen's (1988, as cited in Pallant, 2013) (Cohen's d = 3.88).

Table 4.13

Descriptive Statistics Posttest of Syntactic Accuracy by DCF & NF

Group	N	Mean	SD	SEM
DCF	30	11.05	.21	.03
NF	30	12.07	.30	.05

Table 4.14

Independent T-test Posttest of Syntactic Accuracy by DCF & NF

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval	
								Lower	Upper
Equal variances assumed	2.66	.10	-14.81	58	.000	-1.01	.06	-1.15	-.88
Equal variances not assumed			-14.81	52.47	.000	-1.01	.06	-1.15	-.87

4.5.2. The Long-Term Effect: Delayed-Post-Test Result.

Next, in order to probe the effect of the DCF on the improvement of the syntactic accuracy of the students' essays in the long term, another independent *t*-test was run to compare the

mean scores of the DCF and NF groups on the delayed-posttest of syntactic accuracy in writing.

It should be noted that the assumption of homogeneity of variances was not met (Levene's F = 4.97, p = .03 < .05). That is why the second row of Table 4.16, "Equal variances not assumed" is reported. Based on the results displayed in Tables 4.15 and 4.16, it could be concluded that on average, the DCF group (M = 10.97) outperformed the NF group (M = 11.96). This difference was significant (t (51.28) = -13.24, p < .05); the *p*-value for this t was .000 (Sig (2-tailed) = .000 < .05); however, the magnitude of the differences in the means (mean difference = -.98, 95% CI: -1.13 to -.83 was below medium based on Cohen (1988, as cited in Pallant, 2013) (Cohen's d = 3.69).

Table 4.15

Descriptive Statistics Delayed-Posttest of Syntactic Accuracy by DCF & NF

Group	N	Mean	SD	SEM
DCF	30	10.97	.23	.04
NF	30	11.96	.33	.06

Table 4.16

Independent T-test Delayed-Posttest of Syntactic Accuracy by DCF & NF

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval	
								Lower	Upper
Equal variances assumed	4.97	.03	-13.24	58	.000	-.98	.07	-1.13	-.83
Equal variances not assumed			-13.24	51.28	.000	-.98	.07	-1.13	-.83

4.6. Answer to the Second Research Question

The second research question explored whether there is any significant difference between the effect of ME and NF on developing learners' syntactic accuracy in writing in the short and long term.

4.6.1. The Short-Term Effect: Posttest Result.

An independent *t*-test was run to compare the mean scores of the ME and NF groups on the posttest of syntactic accuracy in writing in order to probe the effect of the ME on the improvement of the syntactic accuracy of the students' essays shortly after the administration of the treatment.

It should be mentioned that the assumption of homogeneity of variances was met (Levene's F = 1.13, p = .29 > .05). That is why the first row of Table 4.18, "Equal variances assumed" is reported. Based on the results displayed in Tables 4.17 and 4.18, it could be concluded that on average, the ME group (M = 10.99) outperformed the NF group (M = 12.07). This difference was significant (t (58) = -15.09, p < .05); the p-value for this t was .000 (Sig (2-tailed) = .000 < .05); however, the magnitude of the differences in the means (mean difference = -1.07, 95% CI: -1.22 to -.93 was below medium based on Cohen (1988, as cited in Pallant, 2013) (Cohen's d = 3.96).

Table 4.17

Descriptive Statistics Posttest of Syntactic Accuracy by ME & NF

Group	N	Mean	SD	SEM
ME	30	10.99	.24	.04
NF	30	12.07	.30	.05

Table 4.18

Independent T-test Posttest of Syntactic Accuracy by ME & NF

	Levene's Test for Equality of Variances		t-test for Equality of Means					
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval
								Lower Upper

Equal variances assumed	1.13	.29	-15.09	58	.000	-1.07	.07	-1.22	-.93
Equal variances not assumed			-15.09	55.22	.000	-1.07	.07	-1.22	-.93

4.6.2. The Long-Term Effect: Delayed-Posttest Result.

Then, in order to probe the effect of the ME on the improvement of the syntactic accuracy of the students' essays in the long term, another independent *t*-test was run to compare the mean scores of the ME and NF groups on the delayed-posttest of syntactic accuracy in writing.

It should be noted that the assumption of homogeneity of variances was not met (Levene's F = 4.69, $p = .03 < .05$). That is why the second row of Table 4.20, "Equal variances not assumed" was reported. Based on the results displayed in Tables 4.19 and 4.20, it could be concluded that on average, the ME group (M = 11.02) outperformed the NF group (M = 11.96). This difference was significant (t (51.53) = -12.65, $p < .05$); the *p*-value for this t was .000 (Sig (2-tailed) = .000 < .05); however, the magnitude of the differences in the means (mean difference = -.94, 95% CI: -1.09 to -.79 was below medium based on Cohen's (1988, as cited in Pallant, 2013) (Cohen's d = 3.52).

Table 4.19

Descriptive Statistics Delayed-Posttest of Syntactic Accuracy by ME & NF

Group	N	Mean	SD	SEM
ME	30	11.02	.23	.04
NF	30	11.96	.33	.06

Table 4.20

Independent T-test Delayed-Posttest of Syntactic Accuracy by ME & NF

	Levene's Test for Equality of Variances		t-test for Equality of Means					
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval

								Lower	Upper
Equal variances assumed	4.69	.03	-12.65	58	.000	-.94	.07	-1.09	-.79
Equal variances not assumed			-12.65	51.53	.000	-.94	.07	-1.09	-.79

4.7. Answer to the Third Research Question

The third research question investigated whether there is any significant difference between the effect of DCF and ME on developing learners' syntactic accuracy in writing in the short and long term.

4.7.1. The Short-Term Effect: Posttest Result.

An independent *t*-test was run to compare the mean scores of the DCF and ME groups on the posttest of syntactic accuracy in writing in order to compare the effect of the two types of treatments on the improvement of the syntactic accuracy of the students' essays shortly after the administration of them.

It should be mentioned that the assumption of homogeneity of variances was met (Levene's F = .41, p = .52 > .05). That is why the first row of Table 4.22, "Equal variances assumed" was reported. Based on the results displayed in Tables 4.21 and 4.22, it could be concluded that there was not a significant difference between the DCF and ME (DCF: M = 11.05; ME: M = 10.99). (t (58) = 1.02, p < .05); the p-value for this t was .31 (Sig (2-tailed) = .31 > .05).

Table 4.21

Descriptive Statistics Posttest of Syntactic Accuracy by DCF & ME

Group	N	Mean	SD	SEM
DCF	30	11.05	.21	.03
ME	30	10.99	.24	.04

Table 4.22

Independent T-test Posttest of Syntactic Accuracy by DCF & ME

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval	
								Lower	Upper
Equal variances assumed	.41	.52	1.02	58	.31	.06	.05	-.05	.18
Equal variances not assumed			1.02	57.32	.31	.06	.05	-.05	.18

4.7.2. The Long-Term Effect: Delayed-Posttest Result.

Then, in order to compare the mean scores of the DCF and ME on the improvement of the syntactic accuracy of the students' essays in the long term, another independent *t*-test (delayed-posttest of syntactic accuracy) was run.

It should be noted that the assumption of homogeneity of variances was met (Levene's F = .000, $p = .93 > .05$). That is why the first row of Table 4.24, "Equal variances assumed" was reported. Based on the results displayed in Tables 4.23 and 4.24, it could be concluded that there was not a significant difference between the DCF and ME. (DCF: M = 10.97; ME: M = 11.02). ($t(58) = -.69$, $p < .05$); the *p*-value for this t was .49 (Sig (2-tailed) = .49 > .05).

Table 4.23

Descriptive Statistics Delayed-Posttest of Syntactic Accuracy by DCF and ME

Group	N	Mean	SD	SEM
DCF	30	10.97	.23	.04
ME	30	11.02	.23	.04

Table 4.24

Independent T-test Delayed-Posttest of Syntactic Accuracy by DCF and ME

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval	
								Lower	Upper
Equal variances assumed	.00	.93	-.69	58	.49	-.04	.05	-.16	.07
Equal variances not assumed			-.69	57.99	.49	-.04	.05	-.16	.07

4.8. Answer to the Fourth Research Question

The fourth research question investigated to what extent the comprehensive DCF and ME are effective on developing EFL learners' written syntactic complexity, relative to each other and to NF (i.e., no feedback) in the short and long term.

4.8.1. The Short-Term Effect: Post-Test Result.

A one-way between-groups multivariate analysis of variance (MANOVA) needed to be performed to investigate the differences between groups in the short term (Larson-Hall, 2010; Pallant, 2013; Tabachnick & Fidell, 2013). This study included 10 dependent variables. MANOVA, however, has a number of assumptions that must be met (Pallant, 2013; Tabachnick & Fidell, 2013). As a result, these assumptions were first examined to be sure that they were not violated. The assumption of univariate normality has been already tested and found to be met.

To test for multivariate normality, Mahalanobis distance was calculated (Pallant, 2013; Tabachnick & Fidell, 2013). The maximum value for Mahalanobis distance for the DCF, ME, and NF were 21.79, 26.05, and 22.72, respectively. Pallant (2013) stated when there are 10 dependent variables, the critical value equals 29.59. The maximum value for

Mahalanobis distance for each group was less than the critical value, so it was safely assumed that there were no multivariate outliers.

Next, to check for multicollinearity and check the strength of the correlations among the dependent variables, correlation was run (Pallant, 2013; Tabachnick & Fidell, 2013). There were no correlations up around .8 or .9 (Pallant, 2013; Tabachnick & Fidell, 2013), so there was no reason for concern regarding multicollinearity. The table indicating the results of correlation coefficients for the posttest of syntactic complexity is included in Appendix 14.

Nevertheless, the correlation coefficients of two indices C/S and CP/T were mostly small, which is considered as very small correlation; therefore, these indices could make the data violate the assumption of linearity. Thus, following Pallant's (2013) recommendation, these indices (i.e., C/S and CP/T) were removed from MANOVA and then separate ANOVAs were performed on them.

Then, to assess the assumption of linearity, a matrix of scatterplots were generated between each pair of variables, separately for each group (DCF, ME, and NF). The scatterplots (See Appendix 15) were roughly oval-shaped (Tabachnick & Fidell, 2013) or cigar-shaped (Pallant, 2013), so it can be concluded that the variables were linearly related and the assumption of linearity was met.

Eventually, preliminary assumption testing to check for univariate and multivariate normality, multicollinearity, as well as linearity revealed no serious violations; thus, a one-way between groups multivariate analysis of variance (MANOVA) was performed to investigate the differences in syntactic complexity in three groups of DCF, ME, and NF. Eight dependent variables were used: MLS, MLT, MLC, VP/T, C/T, DC/C, CT/T, and CN/T. The independent variable was the type of WCF.

Based on Table 4.25, the assumption of homogeneity of variance-covariance was not met because the Sig. value was lower than .001. (.000 < .001), but there was a moderate violation, so as Tabachnick and Fidell (2013) recommended, a more stringent alpha level was adopted. For moderate violations, .025 has been recommended as the alpha level (Tabachnick & Fidell, 2013).

Table 4.25

Box's Test of Equality of Covariance Matrices for Posttest of Syntactic Complexity [a]

Box's M	242.032
F	2.918
df1	72
df2	21089.679
Sig.	.000

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

[a]. Design: Intercept + T

According to Table 4.26, all the Sig. values were greater than .05, indicating that the assumption of equality of variance for the variables was met.

Table 4.26

Levene's Test of Equality of Error Variances for Posttest of Syntactic Complexity [a]

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
MLS	.298	2	87	.743
MLT	1.850	2	87	.163
MLC	.002	2	87	.998
VP/T	.313	2	87	.732
C/T	2.824	2	87	.065
DC/C	1.379	2	87	.257
CT/T	2.543	2	87	.084
CN/T	.038	2	87	.963

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + T

As the second row of Table 4.27, which is for the treatment (i.e., types of WCF) indicates, the *p*-value was less than the stringent alpha level (.024 < .025), so there was a significant difference among the three groups on the combined dependent variables, F (16, 160) = 1.901, p = .024; Wilk's Lambda = .706; partial eta squared = .160. Given the mentioned information, there was a significant difference among the three groups.

Table 4.27

Multivariate Tests for Posttest of Syntactic Complexity

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.996	2537.532[b]	8.000	80.000	.000	.996
	Wilks' Lambda	.004	2537.532[b]	8.000	80.000	.000	.996
	Hotelling's Trace	253.753	2537.532[b]	8.000	80.000	.000	.996
	Roy's Largest Root	253.753	2537.532[b]	8.000	80.000	.000	.996
T	Pillai's Trace	.311	1.865	16.000	162.000	.027	.156
	Wilks' Lambda	.706	1.901[b]	16.000	160.000	.024	.160
	Hotelling's Trace	.392	1.936	16.000	158.000	.021	.164
	Roy's Largest Root	.315	3.188[c]	8.000	81.000	.003	.239

[a] Design: Intercept + T
[b] Exact Statistic
[c] The statistic is an upper bound on F that yields a lower bound on the significance level
[d] Computer using alpha = .05

To determine how the dependent variables differed for the independent variable (i.e., WCF), the Tests of Between-Subjects Effects table was consulted. Because we are looking at a number of separate analyses here, it has been suggested that a higher alpha level be set to reduce the chance of a Type I error (i.e., finding a significant result when there isn't really one) (Pallant, 2013; Tabachnick & Fidell, 2013). The most common way of doing this is to apply a Bonferroni adjustment: dividing the original alpha level of .05 by the number of analyses that is intended to do (i.e., the number of dependent variables) (Pallant, 2013); therefore, in this study, there were eight dependent variables to be investigated by MANOVA; therefore, .05 was divided by 8, giving a new alpha level of .006, so the results were considered significant only if the probability value (Sig.) was less than .006. As the row labeled as T (i.e., Treatment) in Table 4.28 indicates, only one of the dependent variables (C/T) recorded a significant value less than .006. (C/T: F (2, 87 = 2.824, p-value = .004 < .006; partial η^2 = .118, which is almost a large effect size).

Table 4.28

Tests of Between-Subjects Effects for Posttest of Syntactic Complexity

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	MLS	15.948[a]	2	7.974	2.122	.126	.047
	MLT	3.013[b]	2	1.507	.431	.651	.010
	MLC	.620[c]	2	.310	.196	.822	.004
	VP/T	.137[d]	2	.068	.591	.556	.013
	C/T	.502[e]	2	.251	5.831	.004	.118
	DC/C	.008[f]	2	.004	.507	.604	.012
	CT/T	.071[g]	2	.035	1.582	.211	.035
	CN/T	.145[h]	2	.072	.428	.653	.010
Intercept	MLS	22818.739	1	22818.739	6072.109	.000	.986
	MLT	19761.569	1	19761.569	5659.877	.000	.985
	MLC	7841.557	1	7841.557	4961.522	.000	.983
	VP/T	444.023	1	444.023	3837.784	.000	.978
	C/T	245.667	1	245.667	5707.352	.000	.985
	DC/C	11.579	1	11.579	1475.695	.000	.944
	CT/T	19.403	1	19.403	868.451	.000	.909
	CN/T	312.321	1	312.321	1845.782	.000	.955
T	MLS	15.948	2	7.974	2.122	.126	.047
	MLT	3.013	2	1.507	.431	.651	.010
	MLC	.620	2	.310	.196	.822	.004
	VP/T	.137	2	.068	.591	.556	.013
	C/T	.502	2	.251	5.831	.004	.118
	DC/C	.008	2	.004	.507	.604	.012
	CT/T	.071	2	.035	1.582	.211	.035
	CN/T	.145	2	.072	.428	.653	.010
Error	MLS	326.942	87	3.758			
	MLT	303.762	87	3.492			
	MLC	137.501	87	1.580			
	VP/T	10.066	87	.116			
	C/T	3.745	87	.043			
	DC/C	.683	87	.008			
	CT/T	1.944	87	.022			
	CN/T	14.721	87	.169			
Total	MLS	23161.629	90				
	MLT	20068.345	90				
	MLC	7979.678	90				
	VP/T	454.225	90				
	C/T	249.913	90				
	DC/C	12.269	90				
	CT/T	21.418	90				
	CN/T	327.187	90				

Corrected Total	MLS	342.890	89
	MLT	306.775	89
	MLC	138.121	89
	VP/T	10.202	89
	C/T	4.247	89
	DC/C	.691	89
	CT/T	2.014	89
	CN/T	14.866	89

[a]. R Squared = .047 (Adjusted R Squared = .025)
[b]. R Squared = .010 (Adjusted R Squared = -.013)
[c]. R Squared = .004 (Adjusted R Squared = -.018)
[d]. R Squared = .013 (Adjusted R Squared = -.009)
[e]. R Squared = .118 (Adjusted R Squared = .098)
[f]. R Squared = .012 (Adjusted R Squared = -.011)
[g]. R Squared = .035 (Adjusted R Squared = .013)
[h]. R Squared = .010 (Adjusted R Squared = -.013)

It was found that the three groups (i.e., DCF, ME, and NF) differed in terms of C/T, but to find out which group had higher values, the Tukey's HSD post-hoc test was used (Abbott, 2011), which is shown below in the Table 4.29. Table 4.29 indicates that the mean scores for C/T were statistically significantly different between DCF and NF (*p*-value = .004 < .006), but not between ME and NF (*p*- value = .549 > .006), or between DCF and ME (*p*-value = .062 > .006).

Table 4.29

Tukey HSD: Multiple Comparisons for Posttest of Syntactic Complexity for C/T

					95% Confidence Interval	
(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
DCF	ME	.122747	.0535686	.062	-.004987	.250480
	NF	.178847*	.0535686	.004	.051113	.306580
ME	DCF	-.122747	.0535686	.062	-.250480	.004987
	NF	.056100	.0535686	.549	-.071633	.183833
NF	DCF	-.178847*	.0535686	.004	-.306580	-.051113
	ME	-.056100	.0535686	.549	-.183833	.071633

Based on observed means.
The error term is Mean Square (Error) = .169.
*. The mean difference is significant at the .05 level.

These differences are visualized by the plot, as shown in Figure 4.1.

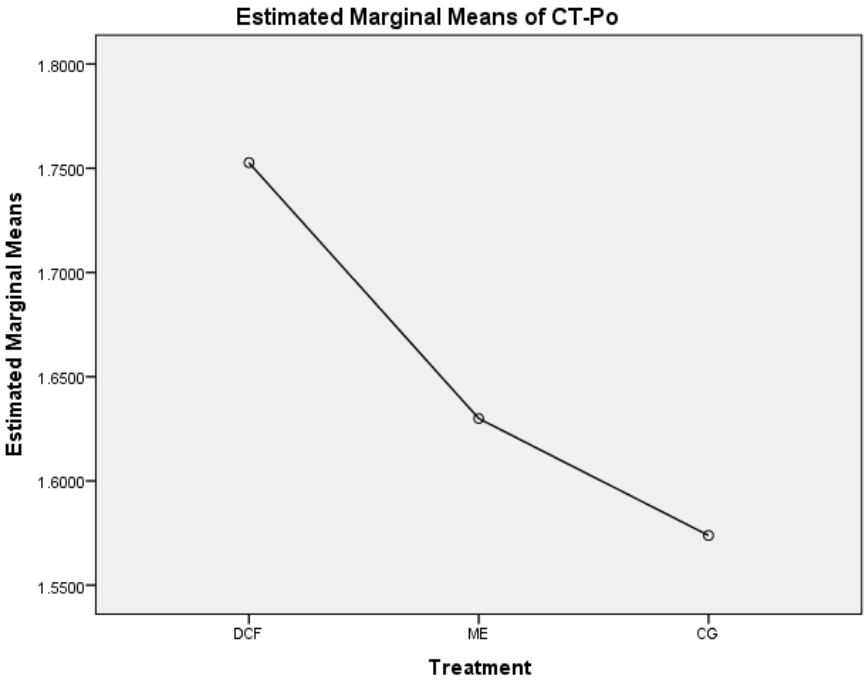


Figure 4.1. Estimated Marginal Means for Posttest of C/T

As it was already mentioned, two indices (C/S and CP/T) were removed from MANOVA test, so two separate ANOVAs were performed on each of them to know whether the three groups were significantly different with regard to these dependent variables.

A one-way between-groups analysis of variance (ANOVA) was conducted to explore whether the three groups (i.e., DCF, ME, NF) were significantly different in the short term with regard to C/S as measured by the posttest. Table 4.30 gives information about each group (number in each group, means, standard deviation, minimum, maximum, etc.).

Table 4.30

Descriptives for Posttest of C/S

Groups	N	Mean	SD	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
DCF	30	2.00	.21	.03	1.92	2.08	1.59	2.47
ME	30	1.76	.26	.04	1.66	1.86	1.13	2.46
NF	30	1.61	.23	.04	1.52	1.69	1.14	2.00
Total	90	1.79	.28	.03	1.73	1.85	1.13	2.47

Table 4.31 indicates that the significance value (Sig.) for Levene's Test is greater than .05 (Sig. = .51), so the assumption of homogeneity of variance has been met.

Table 4.31

Levene's Test for Homogeneity of Variances for posttest of C/S

Levene Statistic	df1	df2	Sig
.66	2	87	.51

As Table 4.32 reveals there was a significant difference between the three groups: F (2, 87) = 20.33, p = .000 (η^2 = .31), which is a large effect size based on the guidelines proposed by Cohen (1988, as cited in Pallant, 2013); it shows 31% of the variance in C/S is explained by the treatment.

Table 4.32

ANOVA for Posttest of C/S

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2.35	2	1.17	20.33	.000
Within Groups	5.04	87	.05		
Total	7.39	89			

To find out which group had the higher scores, this significant ANOVA was followed by Tukey's HSD post-hoc tests (Pallant, 2013). Because two dependent variables were separated from the rest and separate ANOVAs were done on each of them, a more stringent alpha value by using a Bonferroni adjustment was set by dividing the original .05 by 2. The new alpha level became .025. The Table 4.33 shows that mean scores for C/S were statistically significantly different between DCF (M = 2.00, SD = .21) and ME (M = 1.76, SD = .26) (p-value = .001 < .025), and DCF and NF (M = 1.61, SD = .23) (p-value = .000 < .025), but not between ME and NF (p = .037 > .025).

Table 4.33

Tukey HSD: Multiple Comparisons for Posttest of C/S

Multiple Comparisons

Dependent Variable: C/S

Tukey HSD

(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
DCF	ME	.2378067*	.0621578	.001	.089593	.386021
	NF	.3935300*	.0621578	.000	.245316	.541744
ME	DCF	-.2378067*	.0621578	.001	-.386021	-.089593
	NF	.1557233*	.0621578	.037	.007509	.303937
NF	DCF	-.3935300*	.0621578	.000	-.541744	-.245316
	ME	-.1557233*	.0621578	.037	-.303937	-.007509

*. The mean difference is significant at the 0.05 level.

Another one-way between-groups analysis of variance (ANOVA) was conducted to explore whether the three groups (i.e., DCF, ME, NF) were significantly different in the short term with regard to CP/T as measured by the posttest. Table 4.34 gives information about each group (number in each group, means, standard deviation, minimum, maximum, etc.).

Table 4.34

Descriptives for Posttest of CP/T

Groups	N	Mean	SD	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
DCF	30	.54	.12	.02	.49	.59	.29	.81
ME	30	.44	.13	.02	.39	.49	.14	.71
NF	30	.43	.17	.03	.36	.49	.05	.77
Total	90	.47	.15	.01	.44	.50	.05	.81

Table 4.35 indicates that the significance value (Sig.) for Levene's Test is greater than .05 (Sig. = .12), so the assumption of homogeneity of variance has been met.

Table 4.35

Levene's Test for Homogeneity of Variances for Posttest of CP/T

Levene Statistic	df1	df2	Sig
2.094	2	87	.129

As Table 4.36 reveals, there was a significant difference between the three groups: F (2, 87) = 5.220, p = .007 < .025 (η^2 = .10), which is a large effect size based on the guidelines proposed by Cohen (1988, as cited in Pallant, 2013); it shows 10% of the variance in CP/T is explained by the treatment.

Table 4.36

ANOVA for Posttest of CP/T

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.230	2	.115	5.220	.007
Within Groups	1.919	87	.022		
Total	2.149	89			

To find out which group had the higher scores, this significant ANOVA was followed by Tukey's HSD post-hoc tests (Pallant, 2013). The new more stringent alpha level, set by a Bonferroni adjustment is .025. The Table 4.37 shows that mean scores for CP/T were statistically significantly different between DCF (M = .54, SD = .12) and NF (M = .43, SD

= .17) (p-value = .011 < .025), but not between DCF and ME (M = .44, SD = .13) (p-value = .031 > .025), and ME and NF (p = .92 > .025).

Table 4.37

Tukey HSD: Multiple Comparisons for Posttest of CP/T

Multiple Comparisons

Dependent Variable: CP/T
Tukey HSD

(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
DCF	ME	.0990467*	.0383483	.031	.007606	.190487
	NF	.1139933*	.0383483	.011	.022553	.205434
ME	DCF	-.0990467*	.0383483	.031	-.190487	-.007606
	NF	.0149467	.0383483	.920	-.076494	.106387
NF	DCF	-.1139933*	.0383483	.011	-.205434	-.022553
	ME	-.0149467	.0383483	.920	-.106387	.076494

*. The mean difference is significant at the 0.05 level.

4.8.2. The Long-Term Effect: Delayed-Post-Test Result.

Then in order to find out whether there is significant difference between the three groups in the long term, another one-way MANOVA needed to be performed (Larson-Hall, 2010; Pallant, 2013; Tabachnick & Fidell, 2013). As it was already mentioned, this study includes 10 dependent variables. MANOVA, however, has a number of assumptions that must be met (Pallant, 2013; Tabachnick & Fidell, 2013). As a result, these assumptions were first examined to be sure that they were not violated. The assumption of univariate normality has been already tested and found to be met.

To test for multivariate normality, Mahalanobis distance was calculated (Pallant, 2013; Tabachnick & Fidell, 2013). The maximum value for Mahalanobis distance for the DCF, ME, and NF were 22.03, 25.78, and 21.20, respectively. Pallant (2013) stated when there are 10 dependent variables, the critical value equals to 29.59. The maximum value for Mahalanobis distance for each group was less than the critical value, so it was safely assumed that there were no multivariate outliers.

Next, to check for multicollinearity and check the strength of the correlations among the dependent variables, correlation was run (Pallant, 2013; Tabachnick & Fidell, 2013). There were no correlations up around .8 or .9 (Pallant, 2013; Tabachnick & Fidell, 2013), so there was no reason for concern regarding multicollinearity. The table showing the results

of correlation coefficients for the delayed-posttest of syntactic complexity is included in Appendix 16.

Nevertheless, the correlation coefficients of two indices C/S and CP/T were mostly small, which was considered as very small correlation; therefore, these indices could make the data violate the assumption of linearity. Thus, following Pallant's (2013) recommendation, these indices (i.e., C/S and CP/T) were removed from MANOVA and then separate ANOVAs were performed on them.

Then, to assess the assumption of linearity, a matrix of scatterplots were generated between each pair of variables, separately for each group (DCF, ME, and NF). The scatterplots (See Appendix 17) were roughly oval-shaped (Tabachnick & Fidell, 2013) or cigar-shaped (Pallant, 2013), so it was concluded that the variables were linearly related and the assumption of linearity was met.

Eventually, preliminary assumption testing to check for univariate and multivariate normality, multicollinearity, as well as linearity revealed no serious violations; thus, a one-way between groups multivariate analysis of variance (MANOVA) was performed to investigate the differences in syntactic complexity in three groups of DCF, ME, and NF after a two-week interval. Eight dependent variables were used: MLS, MLT, MLC, VP/T, C/T, DC/C, CT/T, and CN/T. The independent variable was the type of WCF.

Based on Table 4.38, the assumption of homogeneity of variance-covariance was not met because the Sig. value was lower than .001. (.000 < .001), but there was a moderate violation, so as Tabachnick and Fidell (2013) recommended, a more stringent alpha level was adopted. For moderate violations, .025 has been recommended as the alpha level.

Table 4.38

Box's Test of Equality of Covariance Matrices for Delayed-Posttest of Syntactic Complexity [a]

Box's M	228.869
F	2.759
df1	72
df2	21089.679
Sig.	.000

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

[a]. Design: Intercept + T

According to Table 4.39, all the Sig. values are greater than .05, indicating that the assumption of equality of variance for the variables has been met.

Table 4.39

Levene's Test of Equality of Error Variances for Delayed-Posttest of Syntactic Complexity[a]

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
MLS	.273	2	87	.762
MLT	1.687	2	87	.191
MLC	.002	2	87	.998
VP/T	1.207	2	87	.304
C/T	2.728	2	87	.071
DC/C	1.381	2	87	.257
CT/T	2.560	2	87	.083
CN/T	.040	2	87	.960

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + T

As the second row of Table 4.40, which is for the treatment (i.e., types of WCF), indicates the *p*-value is less than the stringent alpha level (.002 < .025), so there was a significant difference among the three groups on the combined dependent variables, F (16, 160) = 2.471, p = .002; Wilk's Lambda = .643; partial eta squared = .198. This represents 19.8% of the variance in groups explained by the treatment. Given the mentioned information, there was a significant difference among the three groups.

Table 4.40

Multivariate Tests for Delayed-Posttest of Syntactic Complexity

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.996	2557.430[b]	8.000	80.000	.000	.996
	Wilks' Lambda	.004	2557.430[b]	8.000	80.000	.000	.996
	Hotelling's Trace	255.743	2557.430[b]	8.000	80.000	.000	.996
	Roy's Largest Root	255.743	2557.430[b]	8.000	80.000	.000	.996
T	Pillai's Trace	.380	2.378	16.000	162.000	.003	.190
	Wilks' Lambda	.643	2.471[b]	16.000	160.000	.002	.198
	Hotelling's Trace	.519	2.562	16.000	158.000	.002	.206
	Roy's Largest Root	.435	4.408[c]	8.000	81.000	.000	.303

[a] Design: Intercept + T
[b] Exact Statistic
[c] The statistic is an upper bound by F that yields a lower bound on the significance level
[d] Computer using alpha = .05

To determine how the dependent variables differ for the independent variable (i.e., WCF), the Tests of Between-Subjects Effects table was consulted. Because we are looking at a number of separate analyses here, it has been suggested that a higher alpha level be set to reduce the chance of a Type I error (i.e., finding a significant result when there isn't really one) (Pallant, 2013; Tabachnick & Fidell, 2013). The most common way of doing this is to apply a Bonferroni adjustment. In its simplest form, this involves dividing the original alpha level of .05 by the number of analyses that is intended to do (i.e., the number of dependent variables) (Pallant, 2013); therefore, in this study, there were eight dependent variables to be investigated by MANOVA; therefore, .05 was divided by 8, giving a new alpha level of .006, so the results were considered significant only if the probability value (Sig.) is less than .006. As the row labelled as T (i.e., Treatment) in Table 4.41 indicates, only one of the dependent variables (C/T) recorded a significant value less than .006. (C/T: F (2, 87 = 2.728, p-value = .005 < .006; partial η^2 = .116, which represents 11.6% of the variance in C/T is explained by the treatment.

Table 4.41

Tests of Between-Subjects Effects for Delayed-Posttest of Syntactic Complexity [a]

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	MLS	16.683[a]	2	8.341	2.230	.114	.049
	MLT	2.832[b]	2	1.416	.408	.666	.009
	MLC	.603[c]	2	.302	.191	.827	.004
	VP/T	.611[d]	2	.305	3.837	.025	.081
	C/T	.495[e]	2	.247	5.731	.005	.116
	DC/C	.008[f]	2	.004	.507	.604	.012
	CT/T	.072[g]	2	.036	1.602	.207	.036
	CN/T	.146[h]	2	.073	.432	.651	.010
Intercept	MLS	22779.244	1	22779.244	6090.665	.000	.986
	MLT	19722.244	1	19722.244	5684.904	.000	.985
	MLC	7839.210	1	7839.210	4956.398	.000	.983
	VP/T	401.764	1	401.764	5048.086	.000	.983
	C/T	245.426	1	245.426	5687.874	.000	.985
	DC/C	11.578	1	11.578	1475.991	.000	.944
	CT/T	19.386	1	19.386	867.096	.000	.909
	CN/T	312.244	1	312.244	1846.515	.000	.955
T	MLS	16.683	2	8.341	2.230	.114	.049
	MLT	2.832	2	1.416	.408	.666	.009
	MLC	.603	2	.302	.191	.827	.004
	VP/T	.611	2	.305	3.837	.025	.081
	C/T	.495	2	.247	5.731	.005	.116
	DC/C	.008	2	.004	.507	.604	.012
	CT/T	.072	2	.036	1.602	.207	.036
	CN/T	.146	2	.073	.432	.651	.010
Error	MLS	325.382	87	3.740			
	MLT	301.823	87	3.469			
	MLC	137.602	87	1.582			
	VP/T	6.924	87	.080			
	C/T	3.754	87	.043			
	DC/C	.682	87	.008			
	CT/T	1.945	87	.022			
	CN/T	14.712	87	.169			
Total	MLS	23121.309	90				
	MLT	20026.899	90				
	MLC	7977.416	90				
	VP/T	409.299	90				
	C/T	249.674	90				
	DC/C	12.269	90				
	CT/T	21.403	90				
	CN/T	327.102	90				

Corrected Total	MLS	342.065	89
	MLT	304.655	89
	MLC	138.206	89
	VP/T	7.535	89
	C/T	4.249	89
	DC/C	.690	89
	CT/T	2.017	89
	CN/T	14.858	89

[a]. R Squared = .049 (Adjusted R Squared = .027)
[b]. R Squared = .009 (Adjusted R Squared = -.013)
[c]. R Squared = .004 (Adjusted R Squared = -.019)
[d]. R Squared = .081 (Adjusted R Squared = .060)
[e]. R Squared = .116 (Adjusted R Squared = .096)
[f]. R Squared = .012 (Adjusted R Squared = -.011)
[g]. R Squared = .036 (Adjusted R Squared = .013)
[h]. R Squared = .010 (Adjusted R Squared = -.013)
[i]. Computed using alpha = .05

It was found that the three groups (i.e., DCF, ME, and NF) differed in terms of C/T, but to find out which group had higher values, the Tukey's HSD post-hoc test was used (Abbott, 2011), which is shown below in the Table 4.42. Table 4.42 indicates that the mean scores of C/T were statistically significantly different between DCF and NF (*p*-value = .004 < .006), but not between ME and NF (*p*- value = .542 > .006), or between DCF and ME (*p*-value = .068 > .006).

Table 4.42

Tukey HSD: Multiple Comparisons for Delayed-Posttest of Syntactic Complexity for C/T

(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
DCF	ME	.120987	.0536339	.068	-.006902	.248876
	NF	.177757*	.0536339	.004	.049868	.305646
ME	DCF	-.120987	.0536339	.068	-.248876	.006902
	NF	.056770	.0536339	.542	-.071119	.184659
NF	DCF	-.177757*	.0536339	.004	-.305646	-.049868
	ME	-.056770	.0536339	.542	-.184659	.071119

Based on observed means.
The error term is Mean Square (Error) = .169.
*. The mean difference is significant at the .05 level.

These differences are visualized by the plot, as shown in Figure 4.2.

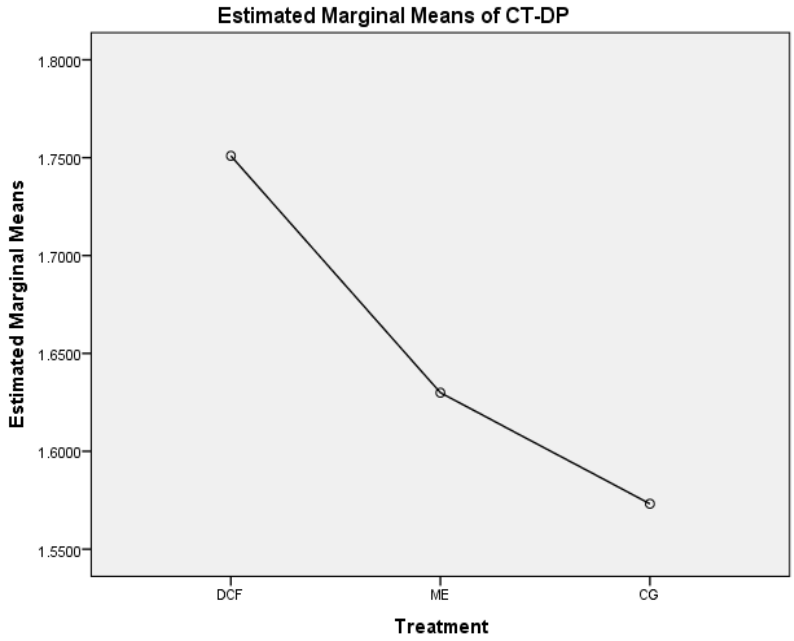


Figure 4.2. Estimated Marginal Means for Delayed-Posttest of C/T

As it was already mentioned, two indices (C/S and CP/T) were removed from MANOVA test, and separate ANOVAs were performed on each of them to know whether the three groups were significantly different with regard to these dependent variables.

A one-way between-groups analysis of variance (ANOVA) was conducted to explore whether the three groups (i.e., DCF, ME, NF) were significantly different in the long term with regard to C/S as measured by the delayed-posttest. Table 4.43 gives information about each group (number in each group, means, standard deviation, minimum, maximum, etc.).

Table 4.43

Descriptives for Delayed-Posttest of C/S

Groups	N	Mean	SD	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
DCF	30	2.00	.20	.03	1.92	2.07	1.59	2.43
ME	30	1.76	.27	.04	1.66	1.86	1.13	2.46
NF	30	1.61	.20	.04	1.54	1.69	1.15	2.00
Total	90	1.79	.27	.02	1.73	1.85	1.13	2.46

Table 4.44 indicates that the significance value (Sig.) for Levene's Test is greater than .05 (Sig. = .32), so the assumption of homogeneity of variance has been met.

Table 4.44

Levene's Test for Homogeneity of Variances for Delayed-Posttest of C/S

Levene Statistic	df1	df2	Sig
1.144	2	87	.323

As Table 4.45 reveals, there was a significant difference between the three groups: F (2, 87) = 20.977, p = .000 (η^2 = .32), which is a large effect size based on the guidelines proposed by Cohen (1988, as cited in Pallant, 2013); it shows 32% of the variance in C/S is explained by the treatment.

Table 4.45

ANOVA for Delayed-Posttest of C/S

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2.233	2	1.117	20.977	.000
Within Groups	4.631	87	.053		
Total	6.864	89			

To find out which group had the higher scores, this significant ANOVA was followed by Tukey's HSD post-hoc tests (Pallant, 2013). Because two dependent variables were separated from the rest and separate ANOVAs were done on each of them, a more stringent alpha value by using a Bonferroni adjustment was set by dividing the original .05 by 2. The new alpha level was .025. The Table 4.46 shows that mean scores for C/S were statistically significantly different between DCF (M = 2.00, SD = .20) and ME (M = 1.76, SD = .27) (p-value = .000 < .025), and DCF and NF (M = 1.61, SD = .20) (p-value = .000 < .025), but not between ME and NF (p = .040 > .025).

Table 4.46

Tukey HSD: Multiple Comparisons for Delayed-Posttest of C/S

Multiple Comparisons

Dependent Variable: C/S

Tukey HSD

(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
DCF	ME	.2348667*	.0595699	.000	.092823	.376910
	NF	.3825467*	.0595699	.000	.240503	.524590
ME	DCF	-.2348667*	.0595699	.000	-.376910	-.092823
	NF	.1476800*	.0595699	.040	.005637	.289723
NF	DCF	-.3825467*	.0595699	.000	-.524590	-.240503
	ME	-.1476800*	.0595699	.040	-.289723	-.005637

*. The mean difference is significant at the 0.05 level.

Another one-way between-groups analysis of variance (ANOVA) was conducted to explore whether the three groups (i.e., DCF, ME, NF) were significantly different in the long term with regard to CP/T as measured by the delayed posttest. Table 4.47 gives information about each group (number in each group, means, standard deviation, minimum, maximum, etc.).

Table 4.47

Descriptives for Delayed-Posttest of CP/T

Groups	N	Mean	SD	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
DCF	30	.52	.17	.03	.45	.58	.24	.89
ME	30	.47	.12	.02	.43	.52	.26	.71
NF	30	.51	.10	.01	.47	.55	.28	.71
Total	90	.50	.13	.01	.47	.53	.24	.89

Table 4.48 indicates that the significance value (Sig.) for Levene's Test was less than .05 (Sig. = .017), so the assumption of homogeneity of variance was violated; thus, Robust Tests of Equality of Means were consulted (Pallant, 2013).

Table 4.48

Levene's Test for Homogeneity of Variances for Delayed Posttest of CP/T

Levene Statistic	df1	df2	Sig
4.251	2	87	.017

As Table 4.49 reveals, there was not a significant difference between the three groups in terms of CP/T.

Table 4.49

Robust Tests of Equality of Means for Delayed-Posttest of CP/T

Robust Tests of Equality of Means

CP/T

	Statistic[a]	df1	df2	Sig.
Welch	.868	2	56.058	.425
Brown-Forsythe	.829	2	72.264	.441

a. Asymptotically F distributed.

4.9. Discussion

4.9.1. Syntactic Accuracy.

Truscott (1996, 1999, 2004, 2007, 2009, 2010) argued that CF should not be used in L2 classrooms due to the fact that correcting the grammatical mistakes would be more likely to prevent than to facilitate accuracy development. However, the findings of the present study clearly indicated that unfocused WCF is effective for improving learners' accuracy; therefore, Truscott's argument that CF has detrimental side effects is not supported by this study.

Some of the early studies did not show promising results for WCF (e.g., Kepner, 1991; Polio et al., 1998; Semke, 1984), and recently Karim and Nassaji (2018) found that only the participants' revisions improved significantly after applying the comprehensive WCF, but the accuracy improvements on new writing text, which was found for DCF and underlining+metalinguistic comprehensive feedback types were non-significant. On the contrary, similar to Robb et al.,'s (1986), Chandler's (2003), Van Beuningan et al.,'s (2008, 2012) as well as Bonilla López et al.,'s (2018)'s findings, this study found that learners who received comprehensive or unfocused WCF made significantly fewer errors in newly produced texts than students who were not provided with CF on their errors.

Considering the goal and scope of the present study, very few studies exist to be exactly compared with the results of the present one because several previous studies which compared the effects of DCF and ME on syntactic accuracy of the written texts explored the focused/selective type of the mentioned WCF types (see e.g., Bitchener, 2008; Bitchener & Knoch, 2008, 2009, 2010b; Mawlawi Diab, 2015; Sheen, 2007; Shintani, Ellis, & Suzuki, 2014; Stefanou & Révész, 2015). The majority of the above-mentioned studies revealed that after receiving the WCF, the students' written accuracy improved significantly in the immediate and delayed post-tests, but the control group did not experience such improvement; further, no significant difference was found between the DCF and ME (Bitchener, 2008; Bitchener & Knoch, 2008, 2009, 2010b; Stefanou & Révész, 2015). However, Sheen (2007) revealed that both direct metalinguistic and direct-only correction groups were significantly better than the control group on the immediate posttests, but the direct metalinguistic group outperformed "the direct-only correction group in the delayed posttests" (p. 255). On the contrary, Shintani et al. (2014), who compared the effects of DCF and ME on university students' accuracy in terms of using

indefinite article and the hypothetical conditional, found that only the accuracy of the hypothetical conditional increased. The DCF was more effective than the ME. The results also indicated that the DCF was more beneficial than ME with regard to errors relevant to complex syntactical structure.

Regarding the present study, it was found that the accuracy of students who received the WCF (i.e., DCF and ME) outperformed those in the control group both in the immediate and delayed post-test and no significant difference was found between the DCF and ME; thus, these results can be similar to the ones found by Bitchener, (2008), Bitchener and Knoch (2008, 2009, 2010b), Stefanou and Révész (2015) as well as Benson and DeKeyser (2018) only in the way that these researchers also found that WCF can enhance the accuracy of the students' written texts. However, the above-mentioned studies investigated the focused type of DCF and ME, not the unfocused or comprehensive feedback like the one conducted in this study.

Different cognitive and socio-cognitive theoretical bases can also be consulted to explain the results of the present study. First, both DCF and ME, in this study, showed positive results, which is in line with the noticing hypothesis (Schmidt, 1990, 2001) as conscious attention to linguistic form facilitates or even is a prerequisite for the development of interlanguage, so CF as a focus-on-form intervention can support the SLA process (DeKeyser, 1994; Han, 2002). It can be stated that by increasing learners' awareness of certain linguistic features, WCF is likely to enable learners to pay attention and be aware of the gaps between their own interlanguage production and the foreign language input (Hulstijn & Schmidt, 1994; Swain, 1991). Subsequently, these noticing operations could have prompted restructuring of learners' developing interlanguage grammar (Gass, 1997; Long, 1996). Moreover, when WCF is provided, learners have enough time – and therefore cognitive resources – to compare their output with the received CF, which raises the likelihood that learners notice gaps in their interlanguage (Polio et al., 1998; Sheen, 2010b). It can be concluded that in the present study, WCF may have helped the learners notice and mentally process the information, and this might have contributed to improvement in written syntactic accuracy also in line with Gass and Mackey (2015).

The results of the current study are also consistent with McLaughlin's (1990) information processing model and Anderson's (1993) ACT (Adaptive Control of Thought) model. Based on these models, information can be processed in either a controlled or automatic way; as a result, learning is a shift from controlled (i.e., declarative knowledge) toward

automatic processing (i.e., automatized procedural knowledge). As a result, it can be argued that intentional learning in this study by means of providing the CF could have played an important role in the controlled phase, and then through 'practice' or 'repeated activation' over time, the language use became automatized. This was also corroborated by several other scholars (see, e.g., Bitchener & Ferris, 2012; DeKeyser, 1997, 2001; Hulstijn, 1995; Schmidt, 1990, 1994, 1995; Swain, 1985; Swain & Lapkin, 1995).

Furthermore, the findings can be supported by the interaction approach to SLA because it is now commonly accepted within the SLA literature that there is a strong connection between the oral or written interaction and learning (Bitchener & Ferris, 2012; Gass & Mackey, 2015; Mackey, 2012; Spada & Lightbown, 2010). It can be explained that in the present study, learning occurred through the participants' exposure to language, production of language and feedback on that production. Socio-cultural theory of human mental processing, based on Vygotsky also assumes that all cognitive development, including language development, takes place as a result of social interactions, especially when learners can collaborate and interact with more knowledgeable speakers of the L2 (e.g., teachers) (Bitchener & Ferris, 2012). Experts in this field (Lantolf & Thorne, 2007; Lantolf et al., 2015; Mackey, 2012) suggested that L2 learners can reach higher levels of linguistic knowledge when they receive proper scaffolding; and learners, with the help of other regulation (e.g., provided by teachers) can finally be self-regulated (i.e., able to use the L2 autonomously). This other regulators may utilize various strategies; one of which is the CF (Lantolf & Thorne, 2007). In short, the WCF types utilized in this study have served the purpose of the research well and the learners received proper scaffolding, so their written syntactic accuracy improved.

4.9.2. Syntactic Complexity.

This study investigated the effects of two different types of comprehensive WCF strategies: DCF and ME on L2 learners' written syntactic complexity, in the short and long term. It was revealed that the DCF significantly outperformed the NF in the C/T, CP/T, and C/S of the texts. DCF also significantly outperformed the ME in the C/S of the text. The ME was not significantly different from the NF. After the two-week interval, the results of the delayed-posttest indicated the DCF still significantly outperformed the NF in terms of the C/T and C/S. DCF also significantly outperformed the ME in C/S, but positive effect of DCF on CP/T was not durable over a two-week period.

Very few studies, which have already investigated the effect of WCF strategies on syntactic complexity in L2 writing found "conflicting findings" (Bitchener & Ferris, 2012, p. 26). For example, Robb et al., (1986)'s study proved that indirect CF was advantageous for the complexity of writing, but later, Sheppard (1992) reported a negative effect. Chandler (2003), using holistic ratings, found no effect on the complexity of L2 students' writing. Van Beuningen et al., (2012) reported that both direct and indirect unfocused WCF did not lead to simplified writing. Finally, more recently, Fazilatfar et al., (2014) found that unfocused WCF had a positive effect on the length of production at MLS, and a dependent clause ratio (DC/C). As for the results of the present study, they are not comparable with the above-mentioned ones because of various reasons which will be elaborated in the following paragraphs.

First, some of the previous studies had some shortcomings. For example, Robb et al.'s (1986) study did not have a no-correction control group (Van Beuningen et al., 2012). Sheppard (1992), in addition to not having a no-correction control group, did not check the inter-rater reliability of the coding of the data, which make us cautious about accepting the findings (Ferris, 2003, 2004). Chandler's (2003) using holistic ratings could not prove that the complexity of students' written texts did not alter (Van Beuningen et al., 2012). Moreover, the participants in Robb et al.'s, (1986) study wrote "narrative compositions" (p. 83); the learners' in Sheppard's (1992) research wrote literary compositions; the students in Chandler's (2003) study wrote autobiographies; the task/genre/content in Van Beuningen et al.'s (2012) study included biology-related topics, all of which differ from the present ones. Genre/task/content/topic can contribute to the development of linguistic complexity (Ortega, 2015; Yoon, 2017). Therefore, the results of the present study cannot be accurately compared with the other mentioned ones.

Furthermore, Lu and Ai (2015) had already indicated that "advanced writers from certain L1 backgrounds may be more prone to produce longer and more complex sentences in English compared to writers from other L1 backgrounds" (p. 11). Given this, skimming through the previously done studies reveals that the participants' L1 in Robb et al.'s, (1986) study, for instance, was Japanese; the 26 participants in Sheppard's (1992) research were from 18 different countries with various L1s; the students in Chandler's (2003) study were from 4 different East Asian counties with different L1s; the participants in Van Beuningen et al.'s (2012) study were multilingual (i.e., with 28 various L1s; and finally, participants' L1 in Fazilatfar et al.'s (2014) study was Persian, whereas the participants' L1 in the

present study was Turkish. Due to the fact that the participants' L1 can be "a moderating variable of L2 syntactic complexity" (Ortega, 2015, p. 84) and that "learners with different L1 backgrounds, even for those at the same or comparable proficiency levels, may not develop in the same ways in all areas." (Lu & Ai, 2015, p. 11), the results of the mentioned studies cannot be truly compared with the present ones.

Furthermore, to include some more points regarding the differences between this study and the other previously mentioned ones, it should be noted that Van Beuningen et al., (2012) worked with participants with limited L2 proficiency, unlike the upper-intermediate participants in the present study. L2 proficiency is "a powerful source of influence that modulates syntactic complexity" (Ortega, 2015, p. 88). This was also proved by Yoon (2017). Additionally, the syntactic complexity measures used by Van Beuningen et al., (2012) differed from the one used in the current study. On the other hand, like the current study, Fazilatfar et al., (2014) used the Lu's (2010) L2 Syntactic and Lexical Complexity Analyzer to assess the syntactic complexity; they only calculated MLS and DC/C, however. They revealed a significant gain for both MLS and DC/C in the experimental group. The present study did not calculate DC/C, and unlike Fazilatfar et al., (2014), the current study did not find a significant effect on MLS.

Despite the above-discussed issues so far, this study, like Robb et al., (1986), Van Beuningen et al., (2012), and Fazilatfar et al., (2014), does not corroborate Truscott's (2004, 2007) speculation that if learners were faced with CF, they would be more likely to avoid using complex structures. With regard to the effect of WCF on L2 learners' written syntactic complexity, the findings of this study also does not conform to the Skehan's (1998) limited capacity models of attention, with which the Truscott's (2004, 2007) claim is consistent because in this study, in both the short and long term, DCF significantly outperformed the NF in the C/T and C/S of the texts; DCF also significantly outperformed the ME in the C/S of the text; and although the ME did not cause a significant positive effect, it did not cause a negative effect either.

As the concluding remark, considering the mentioned points, even though the current study has made an important contribution to the question of whether unfocused WCF can facilitate the development of written syntactic complexity, it is in need of replication before firm conclusions can be reached. Replication of this study seems to be critical to the resolving of doubts about the role of comprehensive WCF in improvement of syntactic complexity. Without a replication, "conclusions will continue to be elusive because each

time different variables or combinations of variables are investigated, the potential exists for such variables to mediate the findings in different ways" (Bitchener & Knoch, 2015, p. 406).

CHAPTER V

CONCLUSIONS

5.1. Overview

This chapter basically includes the following issues: (1) the conclusions drawn by the researcher based on the findings of this study, (2) the theoretical and pedagogical implications that can be drawn from the present work based on the findings (3) limitations which were realized during the study, and (4) the suggestions for future line of studies in the same area.

5.2. Summary and Conclusion

This study aimed at investigating the effects of two different types of comprehensive (unfocused) WCF strategies: direct corrective feedback (DCF), and metalinguistic explanation (ME) on L2 learners' both syntactic accuracy and syntactic complexity in writing, not only in the short term but also in the long term. The participants were 90 Turkish EFL learners who enrolled in upper-intermediate essay writing classes of preparatory university courses. The students who were homogeneous in their writing ability were assigned to three groups, namely DCF (n = 30), ME (n = 30), who received feedback on their grammatical errors and a control group, namely 'No Feedback', abbreviatd as NF (n = 30) who was provided feedback only on content, orthography, and organization, but not on grammatical errors. The treatment/control period lasted for 5 weeks. In the present study, 10 out of 14 indices were analyzed. These ten syntactic complexity measures included: Mean length of clause (MLC), mean length of sentence (MLS), mean length of T-unit (MLT), clauses per sentence (C/S), clauses per T-unit (C/T), complex T-units per T-unit (CT/T), dependent clauses per clause (DC/C), coordinate phrases per T-unit (CP/T), complex nominals per T-unit (CN/T), and verb phrases per T-unit (VP/T). Eventually, both experimental groups significantly outperformed the control one in terms of the syntactic accuracy of the written texts both in the short and long term;

and there was not a significant difference in the DCF and ME both in the short and long term. With regard to short-term effect of the comprehensive WCF types on the syntactic complexity, the DCF statistically significantly outperformed the NF in the C/T, CP/T, and C/S of the texts. DCF also statistically significantly outperformed the ME in the C/S of the text. The ME was not significantly different from the NF in terms of the syntactic complexity indices. After the two-week interval, the results of the delayed posttest indicated that the DCF still statistically significantly outperformed the NF in terms of the C/T and C/S. The DCF also still statistically significantly outperformed the ME in C/S, but positive effect of DCF on CP/T was not durable over a two-week period (i.e., the DCF was not significantly different from the NF in terms of the CP/T in the delayed posttest). In the long-term, the ME was not significantly different from the NF in terms of the syntactic complexity indices.

As it was already mentioned in the discussion section, the findings of the study can be supported by several theoretical cognitive perspectives (information processing models and theories, noticing hypothesis, output hypothesis, and interactionist perspective) as well as socio-cultural perspective. Nevertheless, the results of the present study, especially with regard to the syntactic complexity part, are not comparable to the previous studies because of the differences in the design of the research or due to the shortcomings of some of the previous studies on the issue. Given this, even though the current study has made an important contribution to the question of whether unfocused WCF can facilitate the development of not only written syntactic accuracy but also written syntactic complexity, it is in need of replication before firm conclusions can be reached. Replication of this study seems to be critical to the resolving of doubts about the role of comprehensive WCF in improvement of syntactic accuracy and complexity. Without a replication, "conclusions will continue to be elusive because each time different variables or combinations of variables are investigated, the potential exists for such variables to mediate the findings in different ways" (Bitchener & Knoch, 2015, p. 406).

5.3. Theoretical Implications

The present research (in conjunction with other recent WCF studies) runs counter the claim that CF can only result in 'pseudolearning' as stated by Truscott (1996). The fact that learners benefited from the received WCF when writing new texts indicates that correcting the errors has the potential for accuracy development. Importantly, these CF invoked

accuracy gains proved to be durable; the findings of the delayed-posttests in the present study showed that the observed positive effects of WCF were retained at least over a two-week period; thus, they refute Truscott's (1996) idea that CF can only lead to a transient form of L2 knowledge.

Furthermore, scholars who advocated the focused WCF approach have suggested that focusing on a/a couple of specific error type(s) be more effective than comprehensive or unfocused CF (e.g. Bitchener, 2008; Bitchener & Knoch, 2008, 2009, 2010a, 2010b; Ellis et al., 2008; Sheen, 2007). They stated that comprehensive WCF might fail to facilitate SLA due to the fact that L2 learners have a limited processing capacity and requiring them to deal with a vast range of linguistic features simultaneously may lead to a cognitive overload, and thus prohibit feedback processing. Despite this, the findings of the present study do not support the mentioned hypothesis that processing comprehensive WCF is too cognitively demanding; comprehensive WCF has proved to have the potential to produce sustained accuracy improvement. Moreover, based on limited capacity models of attention, focusing learners' attention on accuracy via the provision of WCF can lead to a reduction of the linguistic complexity of their output (e.g. Skehan, 1998; Skehan & Foster, 2001) because accuracy and complexity use a single source of attention; as a result, these two are always in constant competition for attentional resources. Nevertheless, the findings of the present study are not consistent with this assumption; the learners in the experimental groups not only improved the accuracy of their writing without sacrificing the complexity, but the complexity of the written texts in the DCF also revealed improvement with regard to clauses per sentence (C/S), clauses per T-unit (C/T). In brief, it can be suggested that these results support a multiple-resource perspective on attention (e.g. Robinson, 2003, 2005), based on which, accuracy and complexity are not assumed to be in competition because they are thought to be closely connected.

5.4. Pedagogical Implications

The main and foremost implication of the present findings is that both DCF and ME are useful instruments that L2 teachers can employ to help learners improve their written syntactic accuracy, so teachers should be encouraged to provide either DCF or ME, as both have proven empirically to be significantly effective in developing the learners' syntactic accuracy. Nonetheless, for teachers in search of guidance as to which feedback type to

provide in order to improve the syntactic complexity, DCF proved to be more effective with regard to C/S both in the short and long term in comparison to ME.

However, as Ferris, (2004), K. Hyland and Hyland, (2006), as well as Sheen, (2007) stated, a factor which can have an influence on the degree to which the learners benefit from the WCF is the learners' educational level, and it is presumed that learners' educational level can indicate their level of metalinguistic awareness. That is why learners with lower levels of (meta)linguistic competence might be less able to correct their own errors based on the WCF. As a result, to select the DCF or ME, teachers are suggested to consider the learners' educational level and level of their metalinguistic competence.

It should also be noted that this study, from the outset, attempted to retain ecological classroom validity in WCF research by focusing on a range of syntactic errors, and not on only two or three ones in order to avoid framing the WCF as written grammar exercises rather than authentic writing tasks. Consequently, it can be stated that another important pedagogical contribution of the findings of this thesis is that they demonstrated that the *unfocused* WCF has the potential for yielding a learning effect. As a matter of fact, by exploring the efficacy of the unfocused WCF, the current study attempted to reflect real classroom conditions in terms of the type of provided feedback. That comprehensive WCF as an authentic CF methodology (based on some scholars such as, Ferris (2010), Hartshorn et al., (2010) and Storch (2010)) led to improvement both in syntactic accuracy and also some indices of syntactic complexity in writing could thus be considered of great relevance to teachers.

Nevertheless, as Brown (2017) assured, to select the appropriate types of WCF, which can be appropriate, reasonable, feasible and effective, teachers must carefully consider the contextual factors including the "writing program, assignment type, one's own beliefs as a writing teacher, as well as students' perceptions and what resources they bring to a writing classroom" (p. 111). In addition, as Guénette (2007), Storch (2010), and Van Beuningen (2010) argued, the success or failure of the WCF will depend on a collection of variables including the students' proficiency level, one of the learner-internal factors which can intervene the effect of WCF. Further, with regard to the effect of the WCF on syntactic complexity, as Guénette (2007) argued, "the success or failure of corrective feedback will depend on the classroom context, the type of errors students make, their proficiency level, the type of writing they are asked to do, and a collection of other variables" (pp. 51-52).

Consequently, considering the mentioned contextual and learner factors, it can be suggested that comprehensive WCF can be an effective technique in contexts where teachers need to focus on communicating content or improvement of the overall accuracy of their students' writing rather than on language as an object and learning one or a few specific linguistic feature(s) because the unfocused or comprehensive feedback is mainly concerned with the writing as a whole rather than the writing to practice grammar (Bruton, 2009b; Van Beuningen, 2010, 2011). Moreover, the participants in the current study, who benefited from the unfocused WCF, were at upper-intermediate level of English proficiency; Bitchener and Ferris (2012) had already argued that unfocused WCF would be more useful for the learners of such levels. Therefore, teachers are recommended to apply the comprehensive or unfocused WCF if they work with upper-intermediate or advanced learners with a high level of English proficiency.

5.5. Limitations of the Study

The current study is not without its limitations. In the present study it was attempted to ensure the homogeneity of the participants with regard to their grammatical knowledge, syntactic accuracy and syntactic complexity of the written texts prior to the beginning of the study, but during the research and while correcting the papers and giving feedback, it was found that the students varied widely in the breadth of vocabulary knowledge that they had because of the extensive (i.e., outside the classroom) reading done by some of them. This can be seen as a limitation in this study, which might have had an influence on the results, so it is recommended that future studies regarding the same issue consider both syntactic and lexical homogeneity of the participants.

Another limitation, but with regard to the study design, is that success of the participants in revising their texts was not measured, which is "a necessary first step in the pathway to development" (Brown, 2017, p. 112). Although this was not part of the purpose of the present study, the researcher noticed it would be useful if such purpose had been included in the study, especially because of the ME group, that did not receive the exact correct form of their essays.

5.6. Suggestions for Future Research

Following the findings of this research, several suggestions can be made for further investigation. The followings are the considered suggestions:

1. The length of the course may be critical in gaining results; thus, an approximate replication of this study can be conducted through a longer course of instruction using a longitudinal design. “Providing comprehensive written CF repeatedly over several weeks/months mirrors what learners commonly experience in classroom settings and therefore gaining a better understanding of the effectiveness of such repeated feedback is important not only theoretically but also pedagogically” (Bitchener & Knoch, 2015, p. 412).

2. This study can be replicated with a larger sample of participants at different language proficiency levels to compare the results across these levels. Moreover, in this study, only argumentative type of essay was investigated, this type of essay can be compared with another type in another research. As Bitchener and Knoch (2015) stated, “studies comparing the relative effectiveness of a range of writing tasks and genres and of providing learners of different proficiency levels with the same written CF have yet to be reported” (p. 408)

3. Similar studies can be conducted, but in which the interaction of these two feedback types (i.e., comprehensive DCF and ME) be investigated in relation to specific error types.

4. With regard to focused (selective) feedback, statistically significant positive effects have already been found on posttests and delayed posttests as much as 2–10 months later (e.g., Bitchener, 2008; Bitchener & Knoch, 2008); in this study, it was not possible for the researcher to design an interval longer than 2 weeks, so some critics may think it is still less likely for comprehensive feedback to yield empirically robust findings and be pedagogically effective; thus, another study can be done with longer interval to measure the retention of unfocused (comprehensive) feedback over longer time.

5. As some scholars, such as Ferris (2010) stated and the writing instructors acknowledge, some students benefit more from WCF than others because of “a variety of reasons such as motivation, learning style, and metalinguistic background knowledge” (p. 197), so it is suggested that future research which addresses questions and employs designs similar to the present study, considers and controls the social, contextual, and individual differences, such as the examples mentioned by Ferris (2010), which can mediate the effect of the WCF.

6. A qualitative study is crucial to investigate the students’ attitudes towards the purpose and value of the unfocused DCF and ME as well as their expectations in order to first, investigate their attitudinal engagement, which was suggested by Ellis (2010), and second, finding out some information about their individual differences (IDs) via analyzing their

statements in order to understand how they can be helped to do the writing task in a better way (K. Hyland, 2009) because the mediation of the IDs with the feedback was already proved (Ferris, Liu, Sinha, & Senna, 2013; Y. Han & Hyland, 2015; F. Hyland, 2003; Storch & Wigglesworth, 2010). Moreover, to consider the dynamics of the affective engagement with the intervention, the investigation of the learners' attitudes should be done both during the intervention period and after finishing it due to the fact that according to F. Hyland (2003), Storch and Wigglesworth (2010), Ferris et al., (2013), as well as Han and Hyland (2015), the participants' emotional reactions to the provided intervention may alter during the study.

7. Finally, considering Bitchener and Knoch's (2015) recommendation, think-aloud protocols can be collected from the participants in the two experimental groups (i.e., DCF and ME) while both revising texts and composing new texts in order to "provide in-depth information of how learners process WCF and might give an insight into how learning takes place. Findings from such a study could also have direct implications for theories of L2 acquisition" (p. 412).

REFERENCES

Abbott, M. L. (2011). *Understanding educational statistics using Microsoft EXCEL and SPSS*. New Jersey, US: John Wiley & Sons, Inc.

Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of Proximal development. *The Modern Language Journal*, *78*(4), 465–483. https://doi.org/10.1111/j.1540-4781.1994.tb02064.x

Alkhatib, N. I. M. (2015). *Written corrective feedback at a saudi university: English language teachers' beliefs, students' preferences, and teachers' practices*. (Doctoral dissertation). University of Essex. Retrieved from repository.essex.ac.uk/15382/1/THESIS 6Oct..pdf

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum.

Ashwell, T. (2000). Patterns of teacher response to student writing in a multiple-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing*, *9*(3), 227–258. https://doi.org/10.1016/S1060-3743(00)00027-8

Azizi, M., Behjat, F., & Sorahi, M. A. (2014). Effect of metalinguistic teacher corrective feedback on writing performance of Iranian EFL learners. *International Journal of Language and Linguistics*, *2*(6–1), 54–63. Retrieved from article.sciencepublishinggroup.com/pdf/10.11648.j.ijll.s.2014020601.18.pdf

Bailey, R., & Garner, M. (2010). Is the feedback in higher education assessment worth the paper it is written on? Teachers' reflections on their practices. *Teaching in Higher Education*, *15*(2), 187–198. https://doi.org/10.1080/13562511003620019

Balanga, R. A., Fidel, I. V. B., Gumapac, M. V. G. P., Ho, H. T., Tullo, R. M. C., Villaraza, P. M. L., & Vizconde, C. J. (2016). Student beliefs towards written corrective feedback: The case of Filipino high school students. *I-Manager's Journal on English Language Teaching*, *6*(3), 22–39. Retrieved from https://search.proquest.com/docview/1853319071?pq-origsite=gscholar

Benson, S., & DeKeyser, R. (2018). Effects of written corrective feedback and language aptitude on verb tense accuracy. *Language Teaching Research*, 1–25. https://doi.org/10.1177/1362168818770921

Benton, S. L., Sharp, J. M., Corkill, A. J., Downey, R. G., & Khramtsova, I. (1995). Knowledge, interest, and narrative writing. *Journal of Educational Psychology*, *87*(1), 66–79. https://doi.org/10.1037/0022-0663.87.1.66

Bhatia, V. J. (2002). A generic view of academic discourse. In J. Flowerdew (Ed.), *Academic discourse* (pp. 21–39). London, UK.: Pearson Education.

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, *17*(2), 102–118. https://doi.org/10.1016/j.jslw.2007.11.004

Bitchener, J., East, M., & Cartner, H. (2010). *The effectiveness of providing second language (L2) writers with on-line written corrective feedback*. Retrieved from https://akoaotearoa.ac.nz/download/ng/file/group-5/the-effectiveness-of-providing-second-language-writers-with-on-line-written-corrective-feedback.pdf

Bitchener, J., & Ferris, D. R. (2012). *Written corrective feedback in second language acquisition and writing*. New York, NY: Routledge.

Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research*, *12*(3), 409–431. https://doi.org/10.1177/1362168808089924

Bitchener, J., & Knoch, U. (2009). The relative effectiveness of different types of direct written corrective feedback. *System*, *37*(2), 322–329. https://doi.org/10.1016/j.system.2008.12.006

Bitchener, J., & Knoch, U. (2010a). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing*, *19*(4), 207–217. https://doi.org/10.1016/j.jslw.2010.10.002

Bitchener, J., & Knoch, U. (2010b). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics*, *31*(2), 193–214. https://doi.org/10.1093/applin/amp016

Bitchener, J., & Knoch, U. (2015). Written corrective feedback studies: Approximate replication of Bitchener & Knoch (2010a) and Van Beuningen, de Jong & Kuiken (2012). *Language Teaching*, *48*(3), 405–414. https://doi.org/10.1017/S0261444815000130

Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, *14*(3), 191–205. https://doi.org/10.1016/j.jslw.2005.08.001

Bonilla López, M., Van Steendam, E., Speelman, D., & Buyse, K. (2018). The differential effects of comprehensive feedback forms in the second language writing class. *Language Learning*, *68*(3), 813–850. https://doi.org/10.1111/lang.12295

Brown, D. (2017). *The interaction between grammatical knowledge and explicitness in L2 written corrective feedback*. (Doctoral dissertation). Northern Arizona University. Retrieved from https://search.proquest.com/pqdtglobal/docview/1912717325/fulltextPDF/7CF2D9A0BAFE4886PQ/1?accountid=39519

Bruton, A. (2009a). Designing research into the effects of grammar correction in L2 writing: Not so straightforward. *Journal of Second Language Writing*, *18*(2), 136–140. https://doi.org/10.1016/j.jslw.2009.02.005

Bruton, A. (2009b). Improving accuracy is not the only reason for writing, and even if it were ... *System*, *37*(4), 600–613. https://doi.org/10.1016/j.system.2009.09.005

Bruton, A. (2010). Another reply to Truscott on error correction: Improved situated designs over statistics. *System*, *38*(3), 491–498. https://doi.org/10.1016/j.system.2010.07.001

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, *12*(3), 267–296. https://doi.org/10.1016/S1060-3743(03)00038-9

Chen, S., Nassaji, H., & Liu, Q. (2016). EFL learners' perceptions and preferences of written corrective feedback: A case study of university students from Mainland China. *Asian-Pacific Journal of Second and Foreign Language Education*, *1*(5), 1–17. https://doi.org/10.1186/s40862-016-0010-y

Cohen, A. D. (2012). Strategies: The interface of styles, strategies, and motivation on tasks. In S. Mercer, R. Stephen, & M. Williams (Eds.), *Psychology for language learning: Insights from research, theory and practice* (pp. 136–150). Basingstoke, England: Palgrave Macmillan.

Cohen, A. D., & Cavalcanti, M. (1990). Feedback on written compositions: Teacher and student verbal reports. In B. Kroll (Ed.), *Second Language Writing: Research Insights for the Classroom* (pp. 155–177). Cambridge, UK: Cambridge University.

Crossley, S. A., & McNamara, D. S. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing*, *26*, 66–79. https://doi.org/10.1016/j.jslw.2014.09.006

Cumming, A. (1995). Fostering writing expertise in ESL composition instruction: Modeling and evaluation. In D. Belcher & G. Braine (Ed.), *Academic writing in a second language: Essays on research and pedagogy* (pp. 375–397). Norwood, NJ: Ablex.

DeKeyser, R. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed.), (pp. 94–112). New York, NY: Routledge.

DeKeyser, R. M. (1994). Implicit and explicit learning of second language grammar: A pilot study. *TESOL Quarterly*, *28*(1), 188–194. https://doi.org/10.2307/3587210

DeKeyser, R. M. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition, 19*(02), 195–222. https://doi.org/10.1017/S0272263197002040

DeKeyser, R. M. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125–151). Cambridge, UK: Cambridge University.

Diab, R. L. (2005a). EFL university students' preferences for error correction and teacher feedback on writing. *TESL Reporter*, *38*(1), 27–51. Retrieved from https://laur.lau.edu.lb:8443/xmlui/bitstream/handle/10725/2796/EFL.pdf?sequence=1&isAllowed=y

Diab, R. L. (2005b). Teachers' and students' beliefs about responding to ESL writing: A case study. *TESL Canada*, *23*(1), 28–43. https://doi.org/10.18806/tesl.v23i1.76

Doughty, C. J. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. J. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 256–310). Oxford, UK: Blackwell.

Doughty, C. J., & Williams, J. (1998). Pedagogical choices in focus on form. In C. J. Doughty & J. Williams, *Focus on form in classroom second language acquisition* (pp. 197–261). Cambridge, UK: Cambridge University.

Dusick, D. M. (2011). *Writing the delimitations*. Retrieved from http://bold-ed.com/delimitations.htm

Ellis, R. (1995). Interpretation tasks for grammar teaching. *TESOL Quarterly, 29*(1), 87–106. https://doi.org/10.2307/3587806

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford, UK: Oxford University.

Ellis, R. (2005). Principles of instructed language learning. *System, 33*(2), 209–224. https://doi.org/10.1016/j.system.2004.12.006

Ellis, R. (2009). A typology of written corrective feedback types. *English Language Teaching, 63*(2), 97–107. https://doi.org/10.1093/elt/ccn023

Ellis, R. (2010). EPILOGUE: A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition*, *32*(02), 335–349. https://doi.org/10.1017/S0272263109990544

Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, *36*(3), 353–371. https://doi.org/10.1016/j.system.2008.02.001

Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. New York, NY: Routledge.

Elwood, J. A., & Bode, J. (2014). Student preferences vis-à-vis teacher feedback in university EFL writing classes in Japan. *System*, *42*(1), 333–343. https://doi.org/10.1016/j.system.2013.12.023

Evans, N. W., Hartshorn, K. J., McCollum, R. M., & Wolfersberger, M. (2010). Contextualizing corrective feedback in second language writing pedagogy. *Language Teaching Research*, *14*(4), 445–463. https://doi.org/10.1177/1362168810375367

Farah, M. (2014). *The impact of using flipped classroom instruction on the writing performance of twelfth grade female Emirati students in the applied technology high school (ATHS)*. (Master's thesis). The British University in Dubai (BUiD). Retrieved from https://bspace.buid.ac.ae/bitstream/1234/676/1/120088.pdf

Farrokhi, F., & Sattarpour, S. (2012). The effects of direct written corrective feedback on improvement of grammatical accuracy of high-proficient L2 learners. *World Journal of Education*, *2*(2), 49–57. https://doi.org/10.5430/wje.v2n2p49

Fazilatfar, A. M., Fallah, N., Hamavandi, M., & Rostamian, M. (2014). The effect of unfocused written corrective feedback on syntactic and lexical complexity of L2 writing. In *International Conference on Current Trends in ELT* (Vol. 98), (pp. 482–488). Elsevier B.V. https://doi.org/10.1016/j.sbspro.2014.03.443

Ferris, D. R. (1995a). Can advanced ESL students be taught to correct their most serious and frequent errors? *CATESOL Journal*, *8*(1), 41–62. Retrieved from http://www.catesoljournal.org/wp-content/uploads/2016/11/CJ8_ferris.pdf

Ferris, D. R. (1995b). Student reactions to teacher response in multiple-draft-composition classrooms. *TESOL Quarterly*, *29*(1), 33–53. https://doi.org/10.2307/3587804

Ferris, D. R. (1997). The influence of teacher commentary on student revision. *TESOL Quarterly*, *31*(2), 315–339. https://doi.org/10.2307/3588049

Ferris, D. R. (1999). The case for grammar correction in L2 writing classes, a response to Truscott (1996). *Journal of Second Language Writing*, *8*(1), 1–11. https://doi.org/10.1016/S1060-3743(99)80110-6

Ferris, D. R. (2003). *Response to student writing: Implications for second language students*. Mahwah, NJ: Lawrence Erlbaum Associates.

Ferris, D. R. (2004). The "grammar correction" debate in L2 writing: Where are we, and where do we go from here? (and what do we do in the meantime ...?). *Journal of Response to Writing*, *13*(1), 49–62. https://doi.org/10.1016/j.jslw.2004.04.005

Ferris, D. R. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Ed.), *Feedback in second language writing: Contexts and issues* (pp. 81–104). New York, NY: Cambridge University.

Ferris, D. R. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition*, *32*(02), 181–201. https://doi.org/10.1017/S0272263109990490

Ferris, D. R. (2014). Responding to student writing: Teachers' philosophies and practices. *Assessing Writing*, *19*, 6–23. https://doi.org/10.1016/j.asw.2013.09.004

Ferris, D. R., Brown, J., Liu, H., & Stine, M. E. A. (2011). Responding to L2 writers in college writing classes: What teachers say and what they do. *TESOL Quarterly*, *45*(2), 207–234. https://doi.org/10.5054/tq.2011.247706

Ferris, D. R., Liu, H., Sinha, A., & Senna, M. (2013). Written corrective feedback for individual L2 writers. *Journal of Second Language Writing*, *22*(3), 307–329. https://doi.org/10.1016/j.jslw.2012.09.009

Ferris, D. R., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, *10*(3), 161–184. https://doi.org/10.1016/S1060-3743(01)00039-X

Field, A. (2013). *Discovering statistics using SPSS* (4th. ed.). London: Sage.

Frantzen, D. (1995). The effects of grammar supplementation on written accuracy in an intermediate Spanish content course. *The Modern Language Journal*, *79*(3), 329–344. https://doi.org/10.1111/j.1540-4781.1995.tb01108.x

Frear, D. (2010). The effect of focused and unfocused direct written corrective feedback on a new piece of writing. *College English: Issues and Trends*, *3*, 57–71. Retrieved from http://flc.nccu.edu.tw/Conference/3rd/paper.pdf

Frear, D., & Chiu, Y. H. (2015). The effect of focused and unfocused indirect written corrective feedback on EFL learners' accuracy in new pieces of writing. *System*, *53*, 24–34. https://doi.org/10.1016/j.system.2015.06.006

Gass, S. M. (1997). *Input, interaction and the development of second languages*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S. M., & Mackey, A. (2015). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp. 180–206). New York, NY: Routledge.

Gass, S. M., & Varonis, E. (2004). Input, interaction and second language production. *Studies in Second Language Acquisition*, *16*(3), 283–302. https://doi.org/10.1017/S0272263100013097

Geranpayeh, A. (2003). A quick review of the English quick placement test. *Extract from Research Notes*, *12*, 8–10. Retrieved from http://www.lingue.uniss.it/documenti/lingue/what_is_the_QPT.pdf

Gholaminia, I., Gholaminia, A., & Marzban, A. (2014). An investigation of meta-linguistic corrective feedback in writing performance. In *5th World Conference Educational Sciences-WCES 2013* (pp. 316–320). https://doi.org/10.1016/j.sbspro.2014.01.214

Grami, M. (2005). The effect of teachers' written feedback on ESL students' perception: A study in a Saudi ESL university-level context. *Annual Review of Education, Communication and Language Sciences, 2,* 18–41. Retrieved from http://research.ncl.ac.uk/ARECLS/vol2_documents/Grami/grami.htm

Guénette, D. (2007). Is feedback pedagogically correct? Research design issues in studies of feedback on writing. *Journal of Second Language Writing*, *16*(1), 40–53. https://doi.org/10.1016/j.jslw.2007.01.001

Guénette, D. (2012). The pedagogy of error correction: Surviving the written corrective feedback challenge. *TESL CANADA JOURNAL/REVUE TESL DU CANADA*, *30*(1), 117–126. https://doi.org/10.18806/tesl.v30i1.1129

Hajian, L., Farahani, A. A. K., & Shirazi, M. A. (2014). A study of students' and teachers' preferences and attitudes towards correction of classroom written errors in Iranian EFL context. *International Journal of Applied Linguistics & English Literature*, *3*(5), 287–297. https://doi.org/10.7575/aiac.ijalel.v.3n.5p.287

Hamouda, A. (2011). A study of students and teachers' preferences and attitudes towards correction of classroom written errors in Saudi EFL context. *English Language Teaching, 4*(3), 128–141. https://doi.org/10.5539/elt.v4n3p128

Han, Y., & Hyland, F. (2015). Exploring learner engagement with written corrective feedback in a Chinese tertiary EFL classroom. *Journal of Second Language Writing*, *30*, 31–44. https://doi.org/10.1016/j.jslw.2015.08.002

Han, Z. H. (2002). Rethinking the role of corrective feedback in communicative language teaching. *RELC Journal*, *33*(1), 1–34. https://doi.org/10.1177/003368820203300101

Hartshorn, K. J., Evans, N. W., Merrill, P. F., Sudweeks, R. R., Strong-Krause, D., & Anderson, N. J. (2010). Effects of dynamic corrective feedback on ESL writing accuracy. *TESOL Quarterly*, *44*(1), 84–109. https://doi.org/10.5054/tq.2010.213781

Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. Boston, US: Heinle & Heinle.

Havranek, G. (2002). When is corrective feedback most likely to succeed? *International Journal of Educational Research*, *37*(3–4), 255–270. https://doi.org/10.1016/S0883-0355(03)00004-1

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 1–27). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hedgcock, J., & Lefkowitz, N. (1994). Feedback on feedback: Assessing learner receptivity to teacher response in L2 composing. *Journal of Second Language Writing*, *3*(2), 141–163. https://doi.org/10.1016/1060-3743(94)90012-4

Hedgcock, J., & Lefkowitz, N. (1996). Some input on input: Two analyses of student response to expert feedback in L2 writing. *The Modern Language Journal*, *80*(3), 287–308. https://doi.org/10.1111/j.1540-4781.1996.tb01612.x

Higgins, D., Xi, X., Zechner, K., & Williamson, D. (2011). A three-stage approach to the automated scoring of spontaneous spoken responses. *Computer Speech and Language*, *25*(2), 282–306. https://doi.org/10.1016/j.csl.2010.06.001

Horbacauskiene, J., & Kasperaviciene, R. (2015). Learners' preferences towards corrective feedback in writing assignments in tertiary education. *ExELL (Explorations in English Language and Linguistics)*, *3*(2), 70–83. https://doi.org/10.1515/exell-2017-0002

Hulstijn, J. H. (1995). Not all grammar rules are equal: Giving grammar instruction its proper place in foreign language teaching. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 359–386). Honolulu: University of Hawaii.

Hulstijn, J. H., & Schmidt, R. (1994). Guest editors' introduction. *AILA Review*, *11*, 5–10. Retrieved from http://www.aila.info/download/publications/review/AILA11.pdf

Hyland, F. (1998). The impact of teacher written feedback on individual writers. *Journal of Second Language Writing*, *7*(3), 255–286. https://doi.org/10.1016/S1060-3743(98)90017-0

Hyland, F. (2000). ESL writers and Feedback: Giving more autonomy to students. *Language Teaching Research*, *4*(1), 33–54. https://doi.org/10.1177/136216880000400103

Hyland, F. (2003). Focusing on form: Student engagement with teacher feedback. *System*, *31*(2), 217–230. https://doi.org/10.1016/S0346-251X(03)00021-6

Hyland, K. (2009). *Teaching and researching writing* (2nd ed.). Harlow, UK: Pearson Education Limited.

Hyland, K. (2013). Faculty feedback: Perceptions and practices in L2 disciplinary writing. *Journal of Second Language Writing*, *22*(3), 240–253. https://doi.org/10.1016/j.jslw.2013.03.003

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching*, *39*(02), 83. https://doi.org/10.1017/S0261444806003399

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis. *Studies in Second Language Acquisition*, *24*, 541–577. https://doi.org/10.1017/S0272263102004023

Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *Modern Language Journal*, *99*(1), 1–18. https://doi.org/10.1111/modl.12189

Karim, K., & Nassaji, H. (2018). The revision and transfer effects of direct and indirect comprehensive corrective feedback on ESL students' writing. *Language Teaching Research*, 1–21. https://doi.org/10.1177/1362168818802469

Kassim, A., & Ng, L. L. (2014). Investigating the efficacy of focused and unfocused corrective feedback on the accurate use of prepositions in written work. *English Language Teaching*, *7*(2), 119–130. https://doi.org/10.5539/elt.v7n2p119

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 57–71). Mahwah, NJ: Lawrence Erlbaum Associates.

Kepner, C. G. (1991). An experiment in the relationships of types of written feedback to the development of second language writing skills. *Modern Language Journal*, *75*(3), 305–313. https://doi.org/10.1111/j.1540-4781.1991.tb05359.x

Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing*, *21*(4), 390–403. https://doi.org/10.1016/j.jslw.2012.09.003

Kristia, C. V. (2018). *The relative effects of focused and unfocused direct written corrective feedback on the accurate use of English articles in Hong Kong primary ESL contex.* Retrieved from http://libdr1.ied.edu.hk/pubdata/img00/arch00/link/archive/1/instarh/4718.pdf

Kuiken, F., & Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing*, *17*(1), 48–60. https://doi.org/10.1016/j.jslw.2007.08.003

Lalande, J. F. (1982). Reducing composition errors: An experiment. *The Modern Language Journal*, *66*(2), 140–149. https://doi.org/10.1111/j.1540-4781.1982.tb06973.x

Lantolf, J. P., & Thorne, S. L. (2007). Sociocultural theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction*. Mahwah, NJ: Lawrence Erlbaum Associates.

Lantolf, J. P., Thorne, S. L., & Poehner, M. (2015). Sociocultural theory and second language development. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 207–226). New York, NY: Routledge.

Larsen-freeman, D. (2003). *Teaching language: From grammar to grammaring*. Boston, MA: Heinle & Heinle.

Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. New York, NY: Routledge.

Lee, I. (1997). ESL learners' performance in error correction in writing: Some implications for teaching. *System*, *25*(4), 465–477. https://doi.org/10.1016/S0346-251X(97)00045-6

Lee, I. (2004). Error correction in L2 secondary writing classrooms: The case of Hong Kong. *Journal of Second Language Writing*, *13*(4), 285–312. https://doi.org/10.1016/j.jslw.2004.08.001

Lee, I. (2008). Understanding teachers' written feedback practices in Hong Kong secondary classrooms. *Journal of Second Language Writing*, *17*(2), 69–85. https://doi.org/10.1016/j.jslw.2007.10.001

Lee, I. (2009). Ten mismatches between teachers' beliefs and written feedback practice. *ELT Journal*, *63*(1), 13–22. https://doi.org/10.1093/elt/ccn010

Leki, I. (1991). The preferences of ESL students for error correction in college-level writing classes. *Foreign Language Annals*, *24*(3), 203–218. https://doi.org/10.1111/j.1944-9720.1991.tb00464.x

Liu, Q. (2016). *Effectiveness of coded corrective feedback in the development of linguistic accuracy in L2 writing: Impact of error types and learner attitudes*. (Doctoral dissertation). Northern Arizona University. Retrieved from https://search.proquest.com/pqdtglobal/docview/1818535525/fulltextPDF/B0768679FFED4B3BPQ/1?accountid=39519

Liu, Q., & Brown, D. (2015). Methodological synthesis of research on the effectiveness of corrective feedback in L2 writing. *Journal of Second Language Writing*, *30*, 66–81. https://doi.org/10.1016/j.jslw.2015.08.011

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 438–468). San Diego, CA: Academic.

Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. J. Doughty & J. Williams (Eds.), *Focus on form in second language acquisition* (pp. 15–41). Cambridge, UK: Cambridge University.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, *15*(4), 474–496. https://doi.org/10.1075/ijcl.15.4.02lu

Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, *96*(2), 190–208. https://doi.org/10.1111/j.1540-4781.2011.01232_1.x

Lu, X., & Ai, H. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing*, *29*, 16–27. https://doi.org/10.1016/j.jslw.2015.06.003

Lyster, R., Lightbown, P. M., & Spada, N. (1999). A response to Truscott's 'What's wrong with oral grammar correction''.' *Canadian Modern Language Review*, *55*(4), 457–467. https://doi.org/10.3138/cmlr.55.4.457

Mackey, A. (2012). *Input, interaction, and corrective feedback in L2 learning*. Oxford, UK: Oxford University.

Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Mahfoodh, O. H. A., & Pandian, A. (2011). A qualitative case study of EFL students' affective reactions to and perceptions of their teachers' written feedback. *English Language Teaching*, *4*(3), 14–27. https://doi.org/10.5539/elt.v4n3p14

Mawlawi Diab, N. (2015). Effectiveness of written corrective feedback: Does type of error and type of correction matter? *Assessing Writing*, *24*, 16–34. https://doi.org/10.1016/j.asw.2015.02.001

McCutchen, D. (1986). Domain knowledge and linguistic knowledge in the development of writing ability. *Journal of Memory and Language*, *25*(4), 431–444. https://doi.org/10.1016/0749-596X(86)90036-7

McLaughlin, B. (1990). "Conscious" versus "unconscious" learning. *TESOL Quarterly*, *24*(4), 617–634. https://doi.org/10.2307/3587111

Modirkhamene, S., Soleimani, M., & Sadeghi, K. (2017). Selective vs. comprehensive grammar correction in writing pedagogy: Counter evidence to Truscott's view. *Applied Research on English Language*, *6*(2), 193–212. Retrieved from http://are.ui.ac.ir/article_21347.html

Montgomery, J. L., & Baker, W. (2007). Teacher-written feedback: Student perceptions, teacher self-assessment, and actual teacher performance. *Journal of Second Language Writing, 16*(2), 82–99. https://doi.org/10.1016/j.jslw.2007.04.002

Nassaji, H. (2011). Correcting students' written grammatical errors: The effects of negotiated versus nonnegotiated feedback. *Studies in Second Language Learning and Teaching, 1*(3), 315–334. https://doi.org/10.14746/ssllt.2011.1.3.2

Nassaji, H., & Swain, M. (2000). A Vygotskian perspective on corrective feedback in L2: The effect of random versus negotiated help on the learning of English articles. *Language Awareness, 9*(1), 34–51. https://doi.org/10.1080/09658410008667135

Ortega, L. (2015). Syntactic complexity in L2 writing: Progress and expansion. *Journal of Second Language Writing*, *29*, 82–94. https://doi.org/10.1016/j.jslw.2015.06.008

Pallant, J. (2013). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (5th. ed.). Berkshire, England: Open University.

Phakiti, A. (2010). Analysing quantitative data. In B. Paltridge & A. Phakiti (Eds.), *Continuum companion to research methods in applied linguistics* (pp. 39–49). New York, NY: Continuum Companions.

Pienemann, M. (1987). Determining the influence of instruction on L2 speech processing. *Australian Review of Applied Linguistics*, *10*(2), 83–113. https://doi.org/10.1075/aral.10.2.07pie

Pienemann, M. (1989). Is language teachable? Psycholinguistic experiments and hypotheses. *Applied Linguistics*, *10*(1), 52–79. https://doi.org/10.1093/applin/10.1.52

Pienemann, M. (1998). *Language processing and second language acquisition: Processability theory*. Amsterdam: John Benjamins.

Polio, C., Fleck, C., & Leder, N. (1998). "If only I had more time": ESL learners' changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing*, *7*(1), 43–68. https://doi.org/10.1016/S1060-3743(98)90005-4

Polio, C., & Shea, M. C. (2014). An investigation into current measures of linguistic accuracy in second language writing research. *Journal of Second Language Writing*, *26*, 10–27. https://doi.org/10.1016/j.jslw.2014.09.003

Richards, J. C., & Renandya, A. W. (2002). Teaching writing. In J. C. Richards & A. W. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 303–305). Cambridge, UK: Cambridge University.

Richards, J. C., & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics* (4th ed.). Edinburgh Gate, UK: Pearson Education Limited.

Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly*, *20*(1), 83–93. https://doi.org/10.2307/3586390

Robinson, P. (2003). Attention and memory during SLA. In *The handbook of applied linguisticsandbook of second language acquisition* (pp. 631–678). Oxford, UK: Blackwell.

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *IRAL - International Review of Applied Linguistics in Language Teaching*, *43*(1), 1–32. https://doi.org/10.1515/iral.2005.43.1.1

Sachs, R., & Polio, C. (2007). Learners' uses of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition*, *29*(01), 67–100. https://doi.org/10.10170S0272263107070039

Sadeghi, K., Abolfazli Khonb, Z., & Gheitaranzadeh, F. (2013). The effect of type of corrective feedback (direct vs. indirect) on Iranian pre-intermediate EFL learners' writing. *Biannual Journal of Applied Linguisitcs and Discourse Analysis (JALDA)*, *1*(2), 71–86. Retrieved from http://jalda.azaruniv.ac.ir/article_13527_2313ba5a8fe5ee422fc5d44eb99874dc.pdf

Santos, M., López-Serrano, S., & Manchón, R. M. (2010). The differential effect of two types of direct written corrective feedback on noticing and uptake: Reformulation vs. error correction. *International Journal of English Studies*, *10*(1), 131–154. Retrieved from revistas.um.es/ijes/article/download/114011/108001

Sato, R. (2012). The effects of written feedback in the form of recasts. *Journal of Asia TEFL*, *9*(4), 27–50. Retrieved from www.asiatefl.org/main/download_pdf.php?i=80&c=1419299024

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, *13*, 206–226. https://doi.org/10.1017/S0267190500002476

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *Aila Review*, *11*, 11–26. Retrieved from http://www.aila.info/download/publications/review/AILA11.pdf

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129–158. https://doi.org/10.1093/applin/11.2.129

Schmidt, R. W. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. W. Schmidt (Ed.), *Attention and awareness in foreign language learning and teaching* (pp. 1–63). Honolulu, HI: University of Honolulu.

Schmidt, R. W. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge, UK: Cambridge University.

Schoonen, R. (2005). Generalizability of writing scores: An application of structural equation modeling. *Language Testing*, *22*(1), 1–30. https://doi.org/10.1191/0265532205lt295oa

Semke, H. (1984). The effects of the red pen. *Foreign Language Annals*, *17*(3), 195–202. https://doi.org/10.1111/j.1944-9720.1984.tb01727.x

Sheen, R. (2003). Focus on form–a myth in the making? *ELT Journal*, *57*(3), 225–233. https://doi.org/10.1093/elt/57.3.225

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, *41*(2), 255–283. https://doi.org/10.1002/j.1545-7249.2007.tb00059.x

Sheen, Y. (2010a). Differential effects of oral and written corrective feedback in the ESL classroom. *Studies in Second Language Acquisition*, *32*(02), 203–234. https://doi.org/10.1017/S0272263109990507

Sheen, Y. (2010b). Introduction: The role of oral and written corrective feedback in SLA. *Studies in Second Language Acquisition*, *32*(02), 169–179. https://doi.org/10.1017/S0272263109990489

Sheen, Y. (2011). *Corrective feedback, individual differences and second language learning*. New York: Springer.

Sheen, Y., Wright, D., & Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System*, *37*(4), 556–569. https://doi.org/10.1016/j.system.2009.09.002

Sheppard, K. (1992). Two feedback types: Do they make a difference? *RELC Journal*, *23*(1), 285–304. https://doi.org/10.1177/003368829202300107

Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing*, *22*(3), 286–306. https://doi.org/10.1016/j.jslw.2013.03.011

Shintani, N., & Ellis, R. (2015). Does language analytical ability mediate the effect of written feedback on grammatical accuracy in second language writing? *System*, *49*, 110–119. https://doi.org/10.1016/j.system.2015.01.006

Shintani, N., Ellis, R., & Suzuki, W. (2014). Effects of written feedback and revision on learners' accuracy in using two English grammatical structures. *Language Learning*, *64*(1), 103–131. https://doi.org/10.1111/lang.12029

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, UK: Oxford University.

Skehan, P., & Foster, P. (2001). Cognition and tasks. In *Cognition and second language instruction* (pp. 183–205). Cambridge, UK: Cambridge University.

Smith, M. S. (1981). Consciousness raising and the second language learner. *Applied Linguistics*, *2*(2), 159–168. https://doi.org/10.1093/applin/2.2.159

Smith, M. S. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition*, *15*(02), 165–179. https://doi.org/10.1017/S0272263100011943

Sokolic, M. (2003). Writing. In D. Nunan (Ed.), *Practical English language teaching* (pp. 87–108). New York, NY: McGraw-Hill.

Spada, N., & Lightbown, P. M. (2010). Second language acquisitin. In N. Schmitt (Ed.), *An introduction to applied linguistics* (2nd ed.), (pp. 108–123). Abingdon, Oxon: Hodder Education.

Stefanou, C., & Révész, A. (2015). Direct written corrective feedback, learner differences, and the acquisition of second language article use for generic and specific plural reference. *Modern Language Journal*, *99*(2), 263–282. https://doi.org/10.1111/modl.12212

Storch, N. (2010). Critical feedback on written corrective feedback research. *International Journal of English Studies*, *10*(2), 29–46. Retrieved from https://dialnet.unirioja.es/descarga/articulo/3424315.pdf

Storch, N., & Wigglesworth, G. (2010). Learners' processing, uptake, and retension of corrective feedback on writing: Case studies. *Studies in Second Language Acquisition*, *32*, 303–334. https://doi.org/10.1017/S0272263109990532

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 64–81). Cambridge, UK: Cambridge University.

Swain, M. (1991). French immersion and its offshoots: Getting two for one. In B. Freed (Ed.), *Foreign language acquisition: Research and the classroom* (pp. 91–103). Lexington, MA: Heath.

Swain, M. (1995). Three functions of output in second language learning. In G. Gook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 125–144). Oxford: Oxford University.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook on research in second language teaching and learning*, (pp. 471–484). Mahwah, NJ: Lawrence Erlbaum Associates.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, *16*(3), 371–391. https://doi.org/10.1093/applin/16.3.371

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th. ed.). New Jersey, US: Pearson Education, Inc.

Thode, H. J. (2002). *Testing for normality*. New York, NY: Marcel Dekker.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, *46*(2), 327–369. https://doi.org/10.1111/j.1467-1770.1996.tb01238.x

Truscott, J. (1999). The case for "the case for grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing*, *8*(2), 111–122. https://doi.org/10.1016/S1060-3743(99)80124-6

Truscott, J. (2004). Evidence and conjecture on the effects of correction: A response to Chandler. *Journal of Second Language Writing*, *13*(4), 337–343. https://doi.org/10.1016/j.jslw.2004.05.002

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, *16*(4), 255–272. https://doi.org/10.1016/j.jslw.2007.06.003

Truscott, J. (2009). Arguments and appearances: A response to Chandler. *Journal of Second Language Writing*, *18*(1), 59–60. https://doi.org/10.1016/j.jslw.2008.09.001

Truscott, J. (2010). Some thoughts on Anthony Bruton' s critique of the correction debate. *System*, *38*(2), 329–335. https://doi.org/10.1016/j.system.2010.03.014

Truscott, J., & Hsu, A. Y. (2008). Error correction, revision, and learning. *Journal of Second Language Writing*, *17*(4), 292–305. https://doi.org/10.1016/j.jslw.2008.05.003

Van Beuningan, C. G., De Jong, N. H., & Kuiken, F. (2008). The effect of direct and indirect corrective feedback on L2 learners' written accuracy. *ITL - International Journal of Applied Linguistics*, *156*, 279–296. https://doi.org/10.1075/itl.156.24beu

Van Beuningen, C. G. (2010). Corrective feedback in L2 writing: Theoretical perspectives, empirical insights, and future directions. *International Journal of English Studies*, *10*(2), 1–27. Retrieved from http://revistas.um.es/ijes/article/view/119171

Van Beuningen, C. G. (2011). *The effectiveness of comprehensive corrective feedback in second language writing*. Oisterwijk: Boxpress.

Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, *62*(1), 1–41. https://doi.org/10.1111/j.1467-9922.2011.00674.x

VanPatten, B. (2015). Input processing in adult SLA. In B. VanPatten & J. Williams (Eds.), *Theories in second language acqusition: An introduction* (2nd ed.), (pp. 113–134). New York, NY: Routledge.

Weigle, S. C. (2014). Considerations for teaching second language writing. In M. Celce-Murcia, D. M. Rinton, & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (4th. ed.), (pp. 222–237). Boston, MA: Heinle Cengage Learning.

Westmacott, A. (2017). Direct vs. indirect written corrective feedback: Student perceptions. *Íkala, Revista de Lenguaje y Cultura*, *22*(2), 17–32. https://doi.org/10.17533/udea.ikala.v22n01a02

White, L. (1987). Against comprehensible input: The input hypothesis and the development of second-language competence. *Applied Linguistics*, *8*(2), 95–110. https://doi.org/10.1093/applin/8.2.95

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, *21*(4), 321–331. https://doi.org/10.1016/j.jslw.2012.09.007

Yang, W., Lu, X., & Weigle, S. C. (2015). Different topics, different discourse: Relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. *Journal of Second Language Writing*, *28*, 53– 67. https://doi.org/10.1016/j.jslw.2015.02.002

Yoon, H. J. (2017). Linguistic complexity in L2 writing revisited: Issues of topic, proficiency, and construct multidimensionality. *System*, *66*, 130–141. https://doi.org/10.1016/j.system.2017.03.007

Zacharias, N. T. (2007). Teacher and student attitudes toward teacher feedback. *RELC Journal*, *38*(1), 38–52. https://doi.org/10.1177/0033688206076157

APPENDİXES

Appendix 1. Biodata Information

Please complete the following information sheet. The information will be used only as background information for the present research study and will be kept strictly confidential.

1. Name ……

2. Age ……

3. Where and how long have you studied English? (Check the following.)

Elementary School …… years

Junior High school …… years

High school …… years

University …… years

Institutions …… years

4. Have you ever traveled to or lived in an English-speaking country?

a) Where? (Please specify) ……

……

b) For how long? (Please specify) ……

……

Thank you for your cooperation

Appendix 2. Quick Placement Test

Version 1

Oxford University Press and University of Cambridge Local Examinations Syndicate

Time: 30 minutes

Part 1

Questions 1 – 5

• Where can you see these notices?

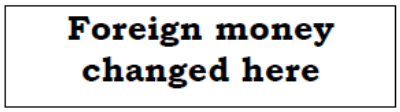
• For questions 1 to 5, mark one letter A, B or C on your Answer Sheet.

1



A in a shop
B in a hotel
C in a taxi

2



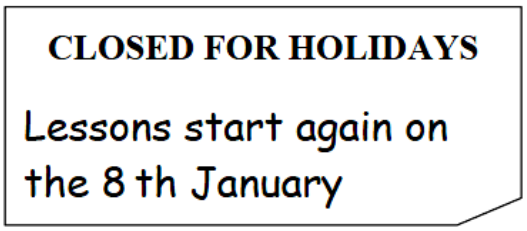
A in a library
B in a bank
C in a police station

3



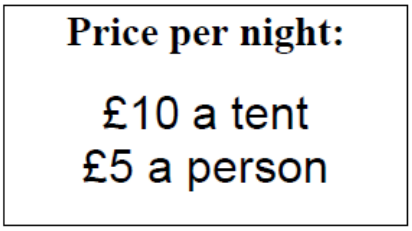
A outside a theatre
B outside a supermarket
C outside a restaurant

4



A at a travel agent's
B at a music school
C at a restaurant

5



A at a cinema
B in a hotel
C on a camp-site

Questions 6 – 10

• In this section you must choose the word which best fits each space in the text below.

• For questions 6 to 10, mark one letter A, B or C on your Answer Sheet.

Scotland

Scotland is the north part of the island of Great Britain. The Atlantic Ocean is on the west and the North Sea on the east. Some people **(6)** ..... Scotland speak a different language called Gaelic. There are **(7)** ..... five million people in Scotland, and Edinburgh is **(8)** ..... most famous city.

Scotland has many mountains; the highest one is called 'Ben Nevis'. In the south of Scotland, there are a lot of sheep. A long time ago, there **(9)** ..... many forests, but now there are only a **(10)** ..... .

Scotland is only a small country, but it is quite beautiful.

6.	A) on	B) in	C) at
7.	A) about	B) between	C) among
8.	A) his	B) your	C) its
9.	A) is	B) were	C) was
10.	A) few	B) little	C) lot

Questions 11 – 20

• In this section you must choose the word which best fits each space in the texts.

• For questions 11 to 20, mark one letter A, B, C or D on your Answer Sheet.

Alice Guy Blaché

Alice Guy Blaché was the first female film director. She first became involved in cinema whilst working for the Gaumont Film Company in the late 1890s. This was a period of great change in the cinema and Alice was the first to use many new inventions, **(11)** ..... sound and colour.

In 1907 Alice **(12)** ..... to New York where she started her own film company. She was **(13)** ..... successful, but, when Hollywood became the centre of the film world, the best days of the independent New York film companies were **(14)** ..... . When Alice died in 1968, hardly anybody **(15)** ..... her name.

11.	A) bringing	B) including	C) containing	D) supporting
12.	A) moved	B) ran	C) entered	D) transported
13.	A) next	B) once	C) immediately	D) recently
14.	A) after	B) down	C) behind	D) over
15.	A) remembered	B) realised	C) reminded	D) repeated

UFOs – do they exist?

UFO is short for 'unidentified flying object'. UFOs are popularly known as flying saucers, **(16)** ..... that is often the **(17)** ..... they are reported to be. The **(18)** ..... "flying saucers" were seen in 1947 by an American pilot, but experts who studied his claim decided it had been a trick of the light.

Even people experienced at watching the sky, **(19)** ..... as pilots, report seeing UFOs. In 1978 a pilot reported a collection of UFOs off the coast of New Zealand. A television **(20)** ..... went up with the pilot and filmed the UFOs. Scientists studying this phenomenon later discovered that in this case they were simply lights on boats out fishing.

16.	A) because	B) therefore	C) although	D) so
17.	A) look	B) shape	C) size	D) type
18.	A) last	B) next	C) first	D) oldest
19.	A) like	B) that	C) so	D) such
20)	A) cameraman	B) director	C) actor	D) announcer

Questions 21 – 40

• In this section you must choose the word or phrase which best completes each sentence.

• For questions 21 to 40, mark one letter A, B, C or D on your Answer Sheet.

21. The teacher encouraged her students ..... to an English pen-friend.

A) should write B) write C) wrote D) to write

22. They spent a lot of time ..... at the pictures in the museum.

A) looking B) for looking C) to look D) to looking

23. Shirley enjoys science lessons, but all her experiments seem to ..... wrong.

A) turn B) come C) end D) go

24. ..... from Michael, all the group arrived on time.

A) Except B) Other C) Besides D) Apart

25. She ..... her neighbour's children for the broken window.

A) accused B) complained C) blamed D) denied

26. As I had missed the history lesson, my friend went ..... the homework with me.

A) by B) after C) over D) on

27. Whether she’s a good actress or not is a ..... of opinion.

A) matter B) subject C) point D) case

28. The decorated roof of the ancient palace was ..... up by four thin columns.

A) built B) carried C) held D) supported

29. Would it ..... you if we came on Thursday?

A) agree B) suit C) like D) fit

30. This form ..... be handed in until the end of the week.

A) doesn’t need B) doesn’t have C) needn’t D) hasn’t got

31. If you make a mistake when you are writing, just ..... it out with your pen.

A) cross B) clear C) do D) wipe

32. Although our opinions on many things ..... , we’re good friends.

A) differ B) oppose C) disagree D) divide

33. This product must be eaten ..... two days of purchase.

A) by B) before C) within D) under

34. The newspaper report contained ..... important information.

A) many B) another C) an D) a lot of

35. Have you considered ..... to London?

A) move B) to move C) to be moving D) moving

36. It can be a good idea for people who lead an active life to increase their ..... of vitamins.

A) upturn B) input C) upkeep D) intake

37. I thought there was a ..... of jealousy in his reaction to my good fortune.

A) piece B) part C) shadow D) touch

38. Why didn’t you ..... that you were feeling ill?

A advise B mention C remark D tell

39. James was not sure exactly where his best interests ..... .

A) stood B) rested C) lay D) centred

40. He’s still getting ..... the shock of losing his job.

A) across B) by C) over D) through

Part 2

Questions 41 – 50

• In this section you must choose the word or phrase which best fits each space in the texts.

• For questions 41 to 50, mark one letter A, B, C or D on your Answer Sheet.

The tallest buildings - SKYSCRAPERS

Nowadays, skyscrapers can be found in most major cities of the world. A building which was many **(41)** …….. high was first called a skyscraper in the United States at the end of the 19th century, and New York has perhaps the **(42)** …….. skyscraper of them all, the Empire State Building. The **(43)** …….. beneath the streets of New York is rock, **(44)** …….. enough to take the heaviest load without sinking, and is therefore well-suited to bearing the **(45)** …….. of tall buildings.

41.	A) stages	B) steps	C) storeys	D) levels
42.	A) first-rate	B) top-class	C) well-built	D) best-known
43.	A) dirt	B) field	C) ground	D) soil
44.	A) hard	B) stiff	C) forceful	D) powerful
45.	A) weight	B) height	C) size	D) scale

SCRABBLE

Scrabble is the world's most popular word game. For its origins, we have to go back to the 1930s in the USA, when Alfred Butts, an architect, found himself out of **(46)** …….. . He decided that there was a **(47)** ……. for a board game based on words and **(48)** ……. To design one. Eventually he made a **(49)** ……. from it, in spite of the fact that his original **(50)** ……. was only three cents a game.

46.	A) earning	B) work	C) income	D) job
47.	A) market	B) purchase	C) commerce	D) sale
48.	A) took up	B) set out	C) made for	D) got round
49.	A) wealth	B) fund	C) cash	D) fortune
50.	A) receipt	B) benefit	C) profit	D) allowance

Questions 51 – 60

• In this section you must choose the word or phrase which best completes each sentence.

• For questions 51 to 60, mark one letter A, B, C or D on your Answer Sheet.

51. Roger's manager ..... to make him stay late if he hadn't finished the work.

A) insisted B) warned C) threatened D) announced

52. By the time he has finished his week's work, John has hardly ..... energy left for the weekend.

A) any B) much C) no D) same

53. As the game ..... to a close, disappointed spectators started to leave.

A) led B) neared C) approached D) drew

54. I don't remember ..... the front door when I left home this morning.

A) to lock B) locking C) locked D) to have locked

55. I ..... to other people borrowing my books: they always forget to return them.

A) disagree B) avoid C) dislike D) object

56. Andrew's attempts to get into the swimming team have not ..... with much success.

A) associated B) concluded C) joined D) met

57. Although Harry had obviously read the newspaper article carefully, he didn't seem to have ..... the main point.

A) grasped B) clutched C) clasped D) gripped

58. A lot of the views put forward in the documentary were open to ..... .

A) enquiry B) query C) question D) wonder

59. The new college ..... for the needs of students with a variety of learning backgrounds.

A) deals B) supplies C) furnishes D) caters

60. I find the times of English meals very strange – I'm not used ..... dinner at 6pm.

A) to have B) to having C) having D) have

Appendix 3. Writing Topics

(Week 1) Pretest – Task 1:

Do you agree or disagree with the following statement?

People should read only those books that are about real events, real people, and established facts.

Use specific reasons and details to support your opinion.

(Week 2) Task 2

Do you agree or disagree with the following statement?

Watching television is bad for children.

Use specific details and examples to support your answer.

(Week 3) Task 3

Do you agree or disagree with the following statement?

Children should begin learning a foreign language as soon as they start school.

Use specific reasons and examples to support your position.

(Week 4) Task 4

Do you agree or disagree with the following statement?

Classmates are a more important influence than parents on a child's success in school.

Use specific reasons and examples to support your answer.

(Week 5) Task 5

Do you agree or disagree with the following statement?

Parents or other adult relatives should make important decisions for their (15 to 18 year-old) teenage children.

Use specific reasons and examples to support your opinion.

(Week 6) Posttest – Task 6

Do you agree or disagree with the following statement?

Most experiences in our lives that seemed difficult at the time become valuable lessons for the future.

Use reasons and specific examples to support your answer.

(Week 8) Delayed-Posttest – Task 7

Do you agree or disagree with the following statement?

Face-to-face communication is better than other types of communication, such as letters, email, or telephone calls.

Use specific reasons and details to support your answer.

Appendix 4. Histograms for QPT

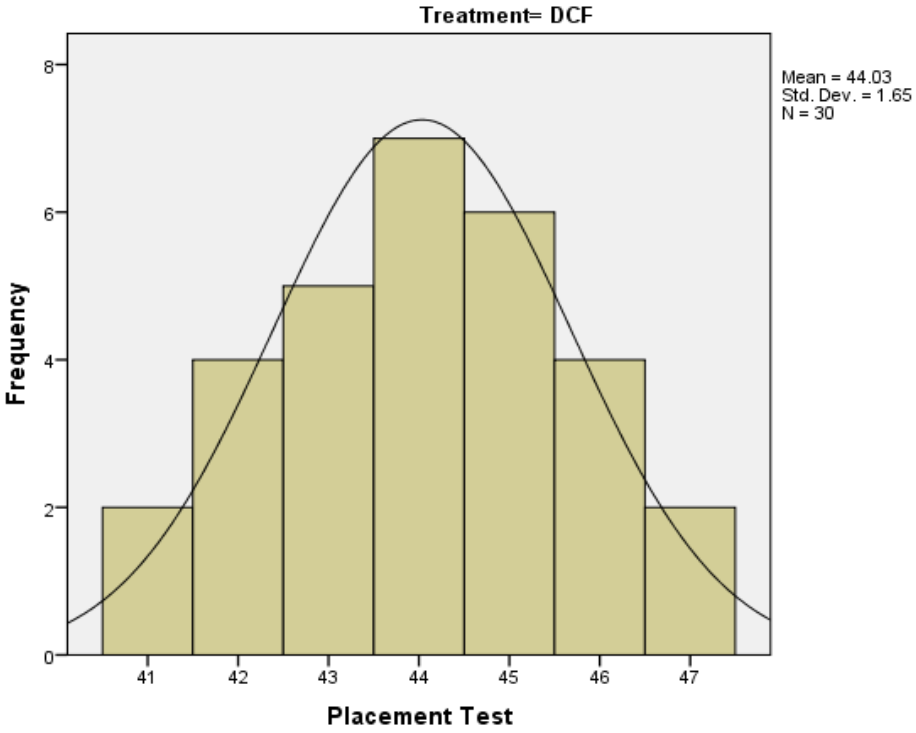


Figure 4.3. QPT in the DCF Group

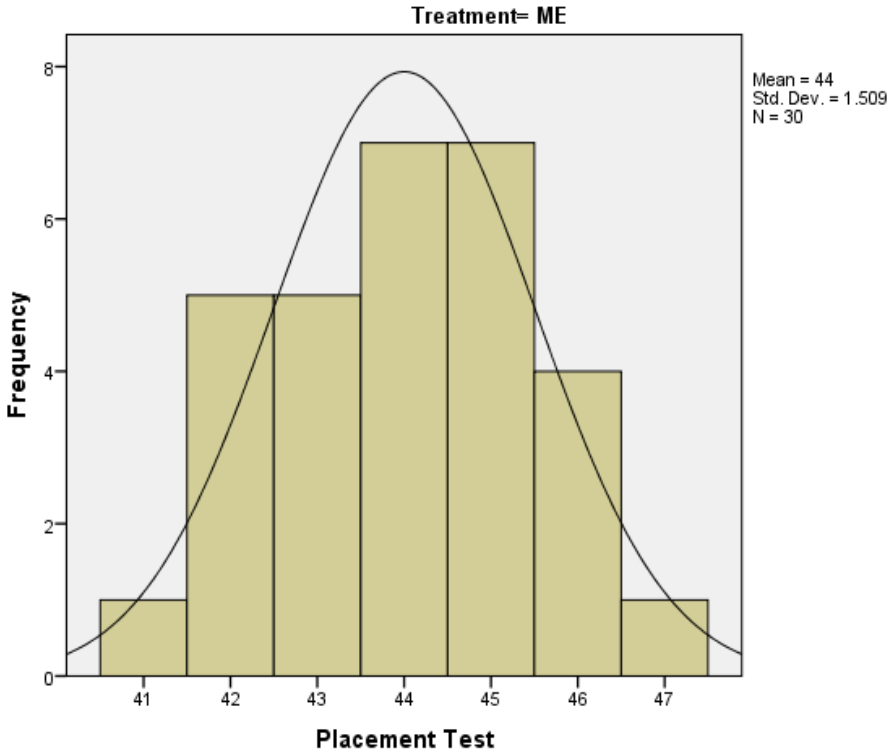


Figure 4.4. QPT in the ME Group

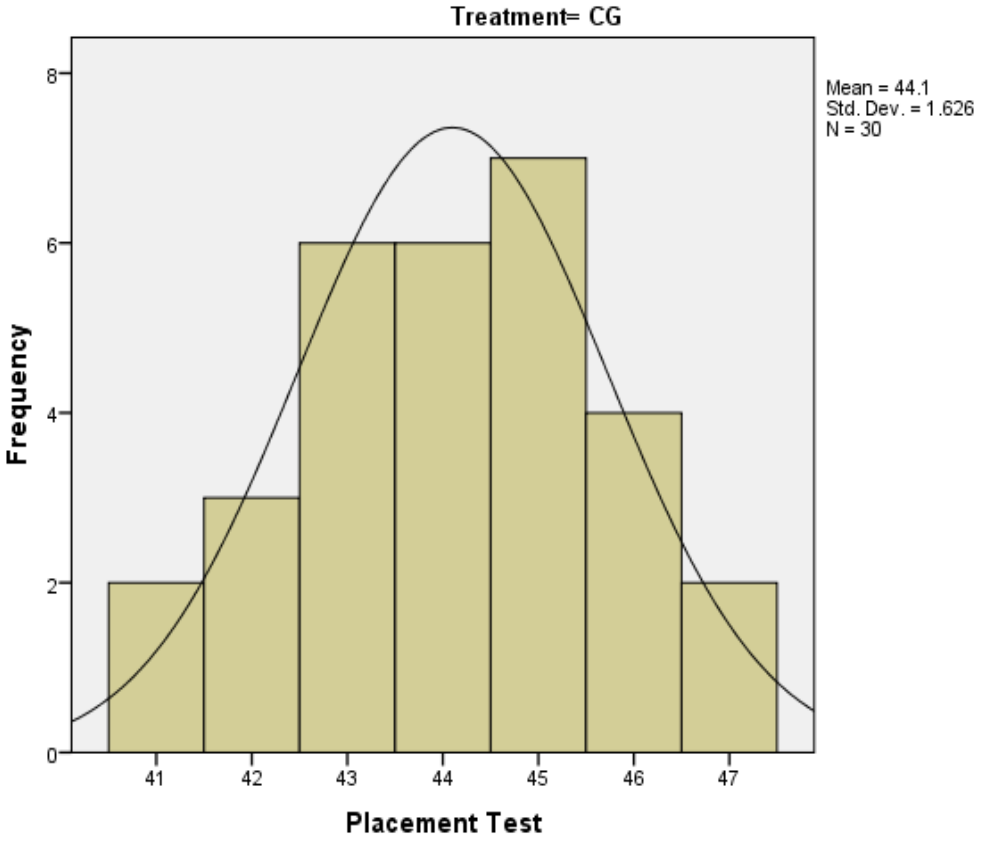


Figure 4.5. QPT in the NF Group

Appendix 5. Histograms for syntactic accuracy measures (pretest, posttest, and delayed-posttest)

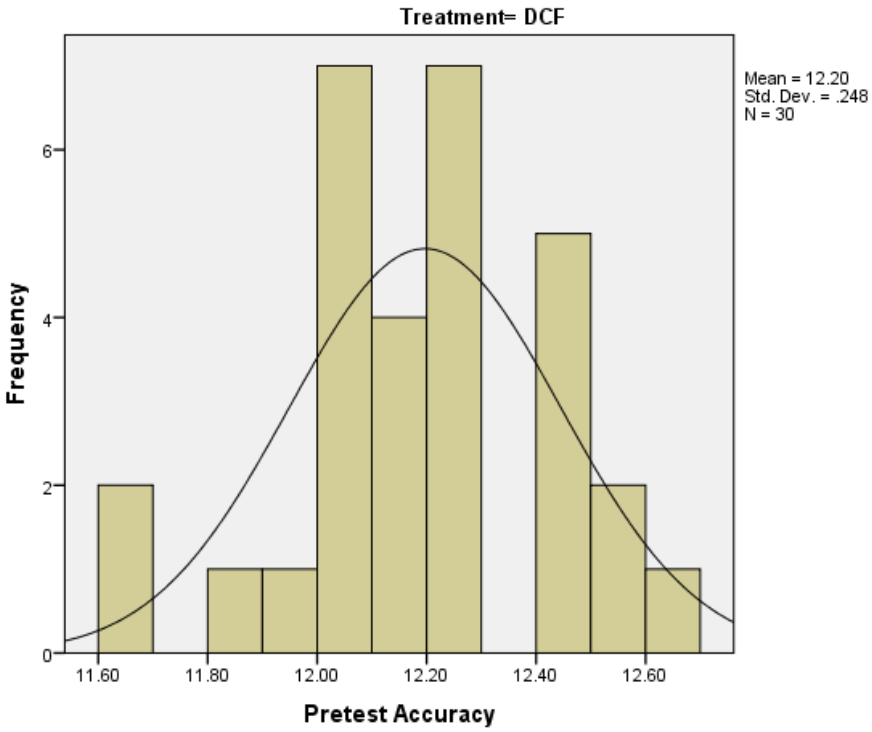


Figure 4.6. Pretest of Syntactic Accuracy in the DCF Group

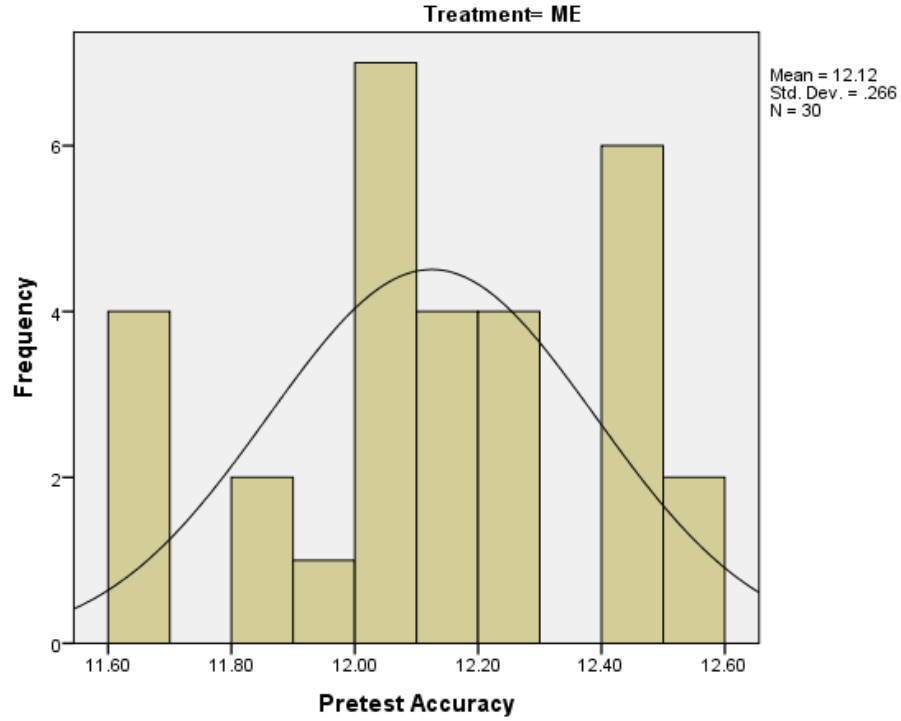


Figure 4.7. Pretest of Syntactic Accuracy in the ME Group

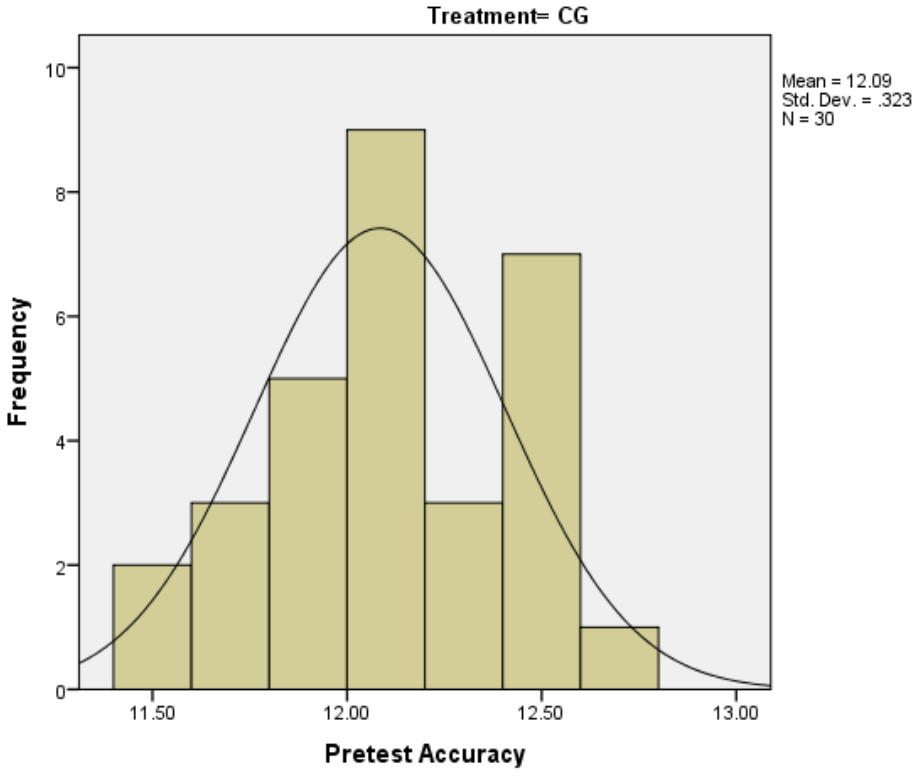


Figure 4.8. Pretest of Syntactic Accuracy in the NF Group

Figure 4.9. Posttest of Syntactic Accuracy in the DCF Group

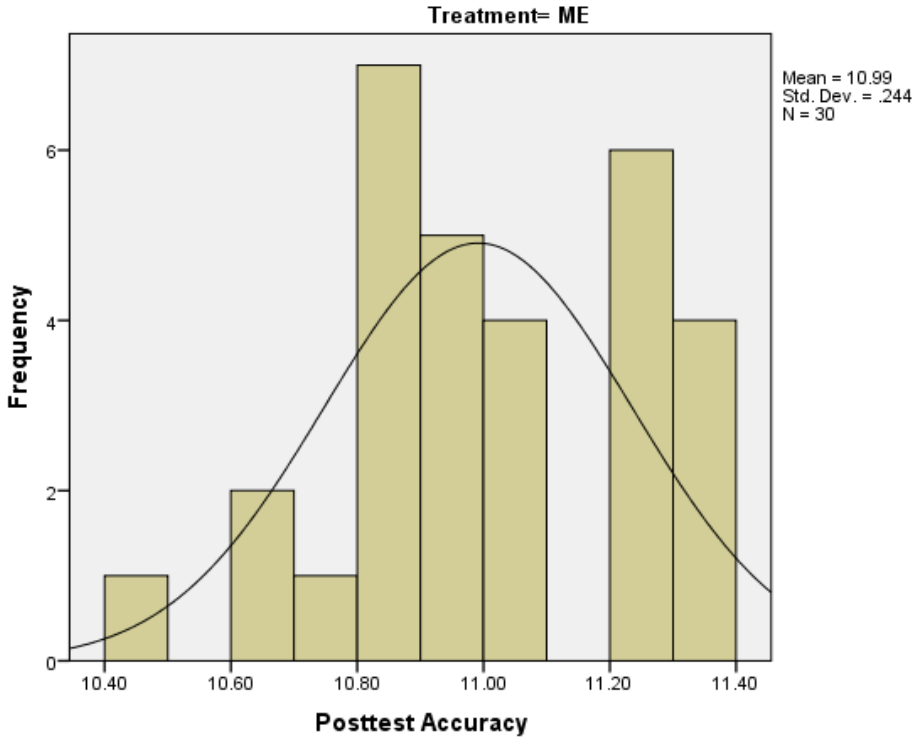


Figure 4.10. Posttest of Syntactic Accuracy in the ME Group

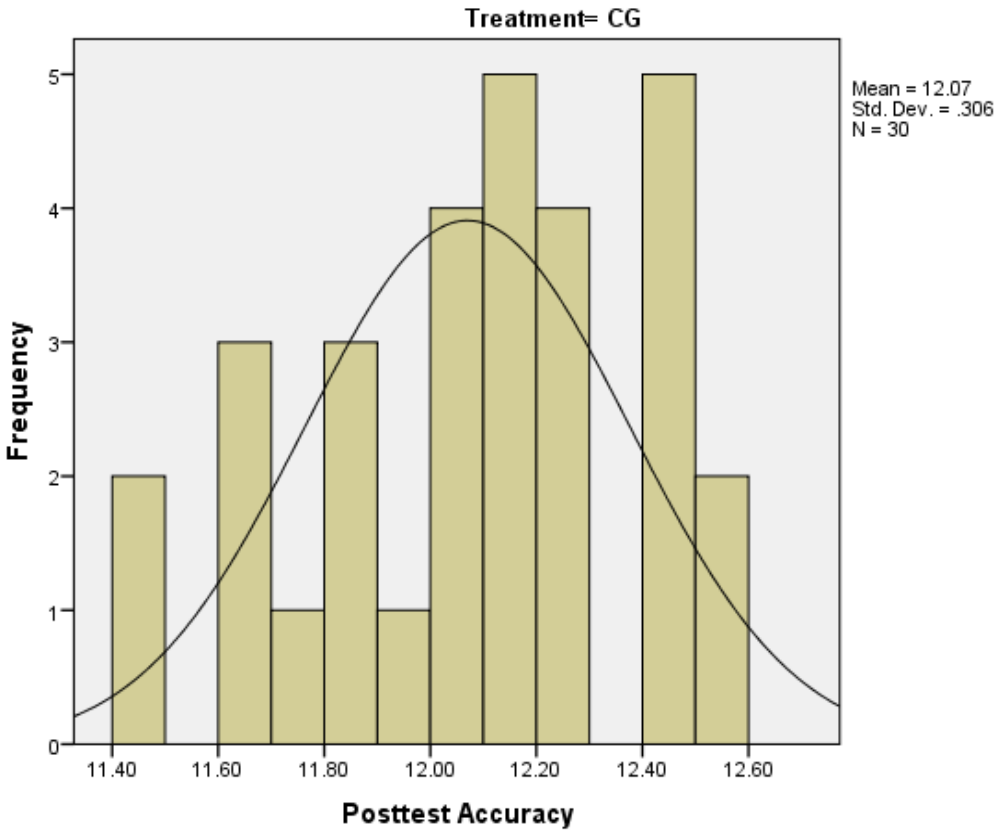


Figure 4.11. Posttest of Syntactic Accuracy in the NF Group

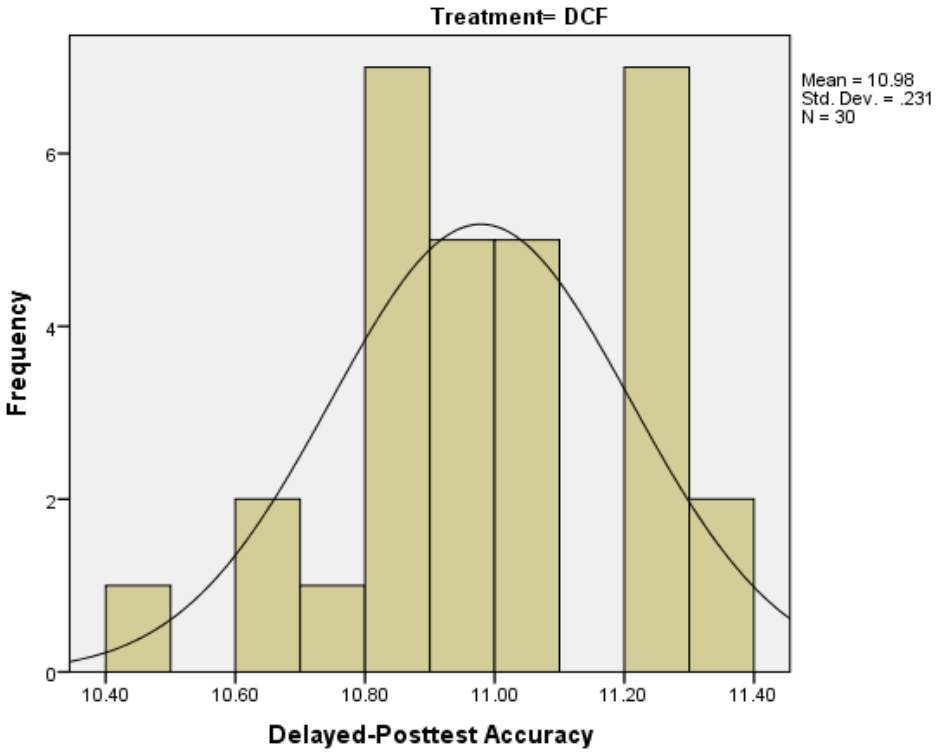


Figure 4.12. Delayed-Posttest of Syntactic Accuracy in the DCF Group

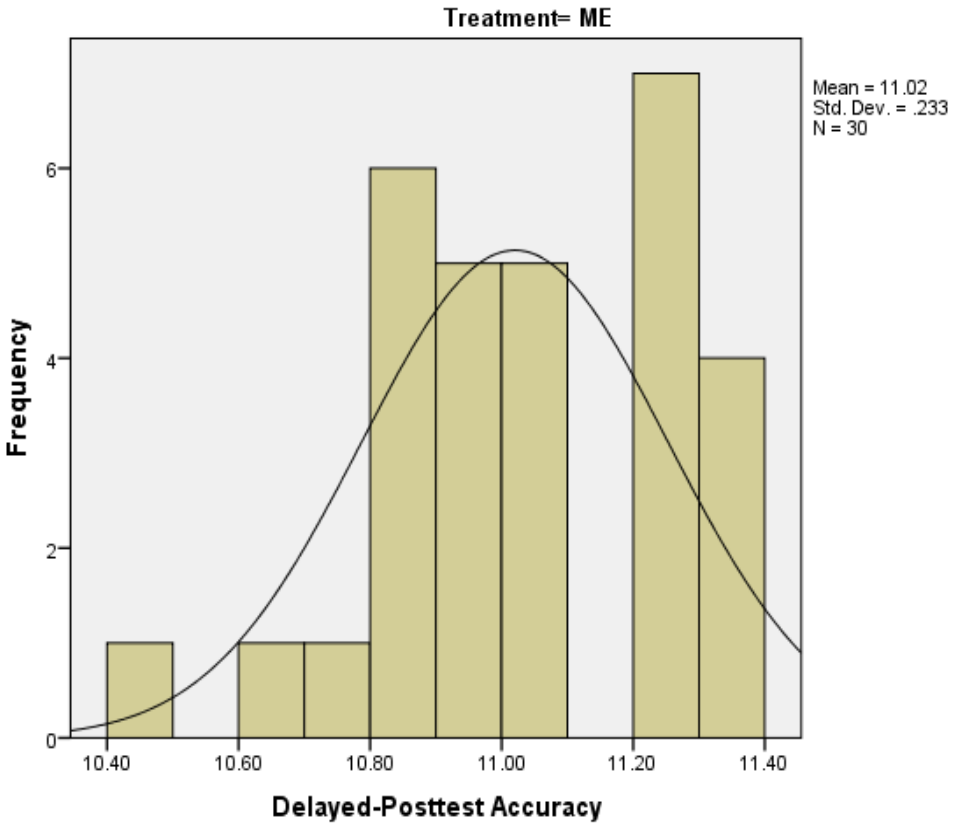


Figure 4.13. Delayed-Posttest of Syntactic Accuracy in the ME Group

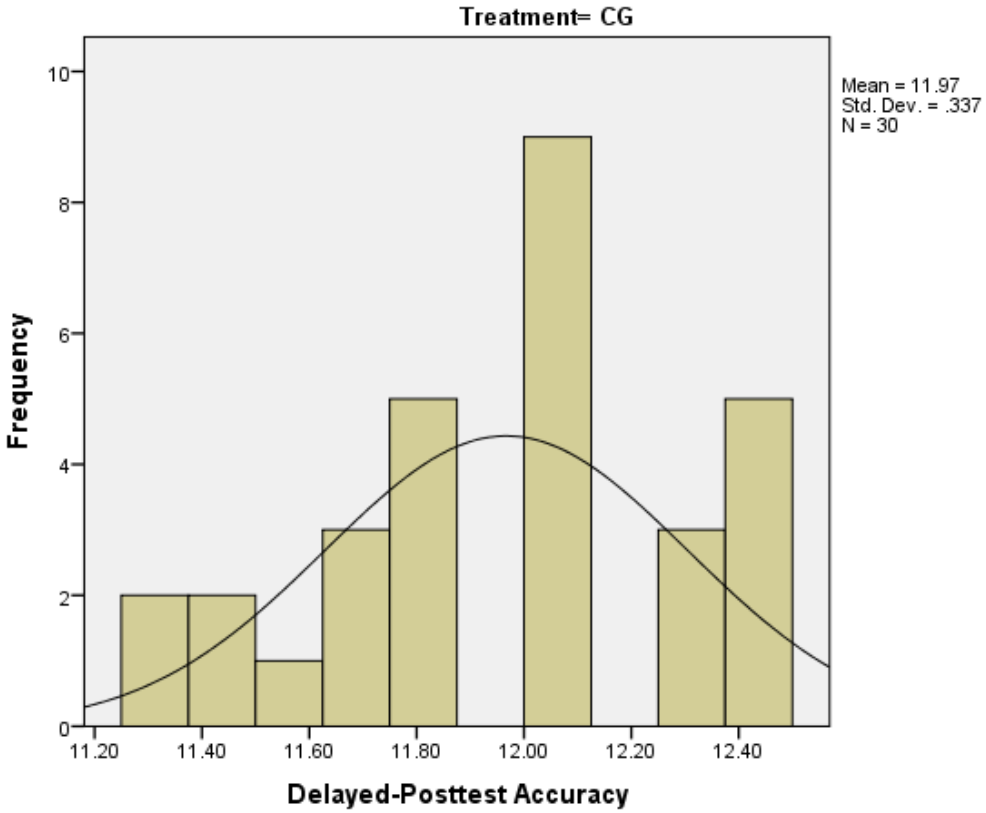


Figure 4.14. Delayed-Posttest of Syntactic Accuracy in the NF Group

Appendix 6. Histograms for Pretest of Syntactic Complexity

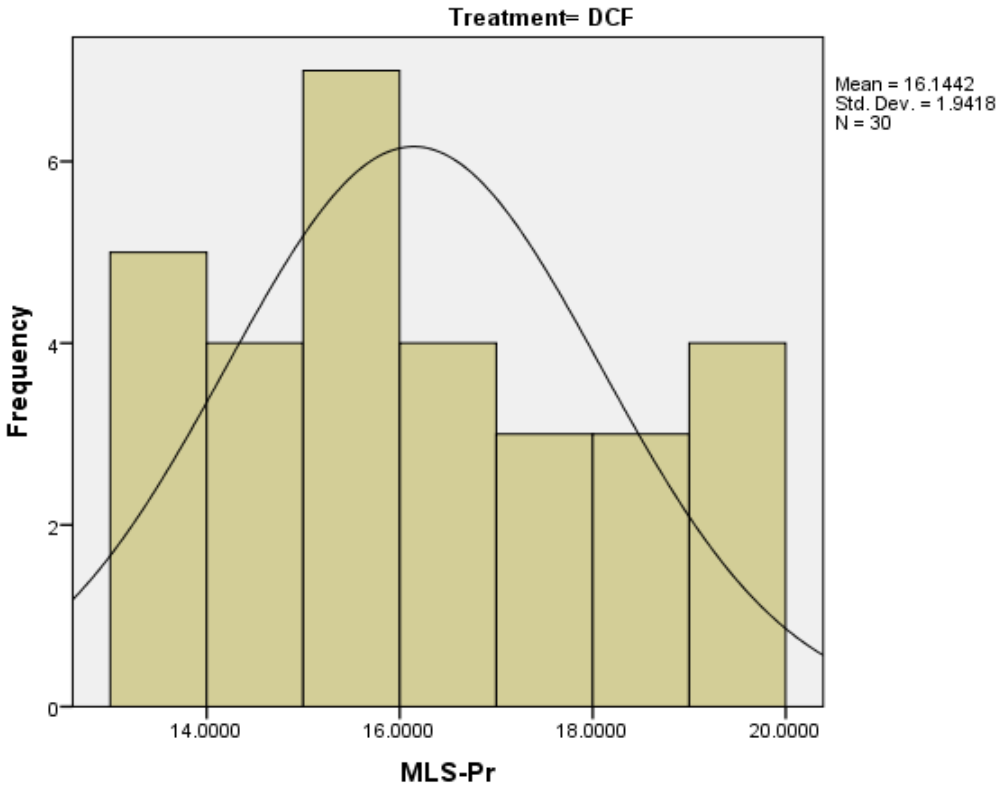


Figure 4.15. Pretest of MLS in the DCF Group

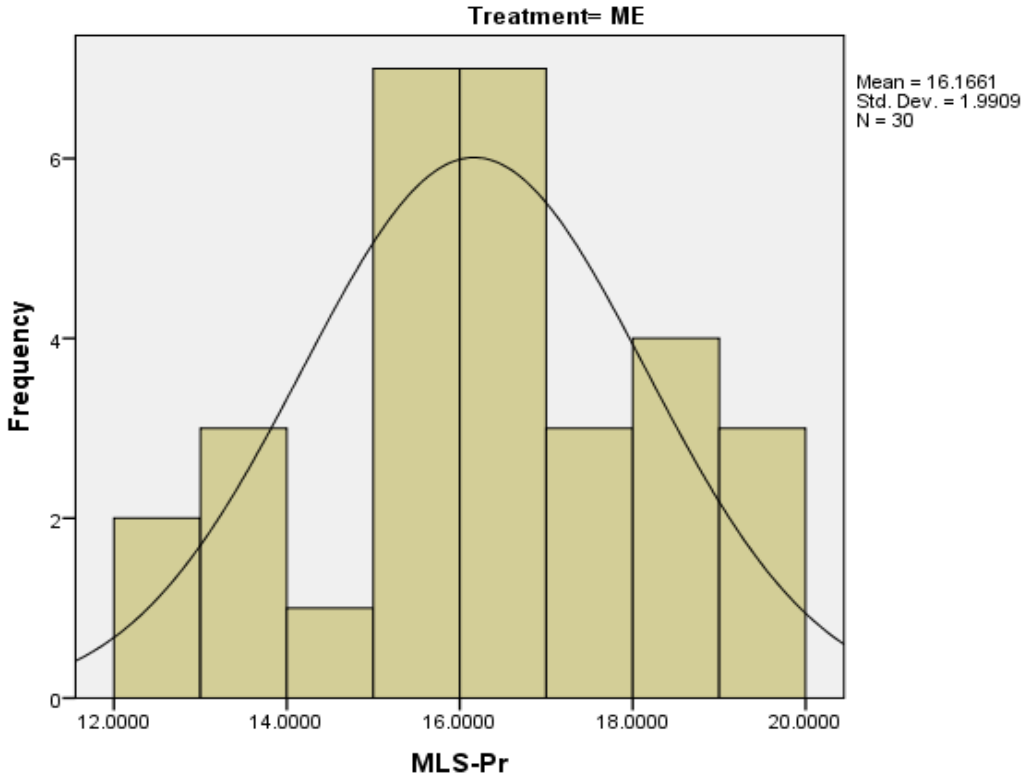


Figure 4.16. Pretest of MLS in the ME Group

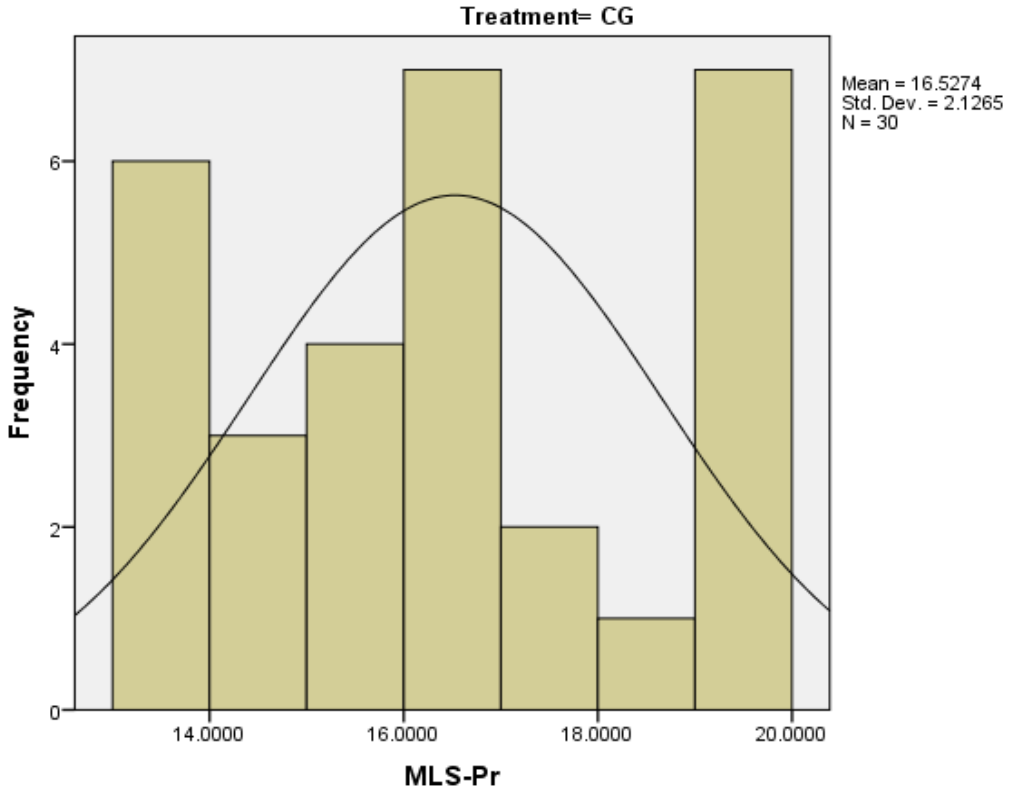


Figure 4.17. Pretest of MLS in the NF Group

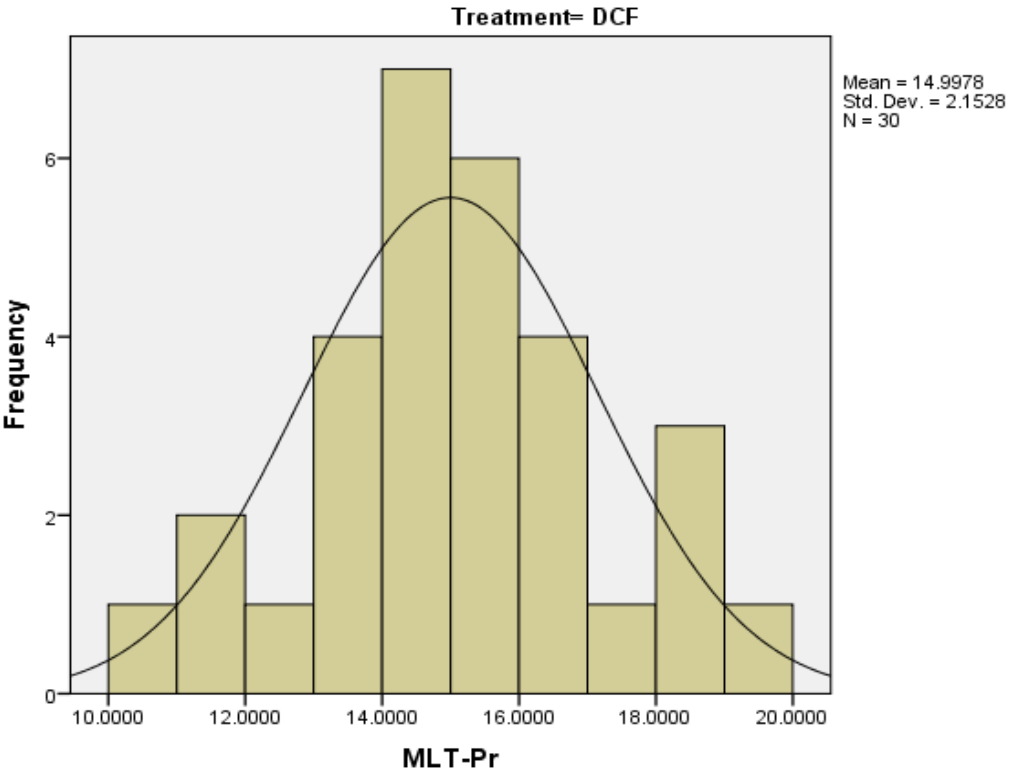


Figure 4.18. Pretest of MLT in the DCF Group

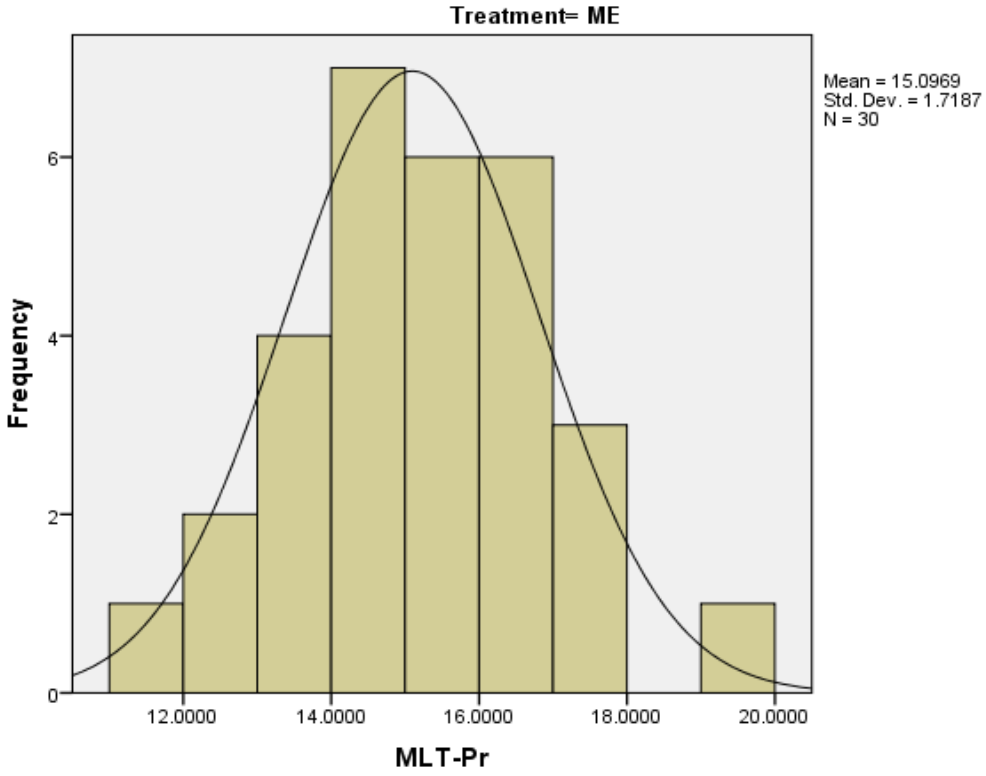


Figure 4.19. Pretest of MLT in the ME Group

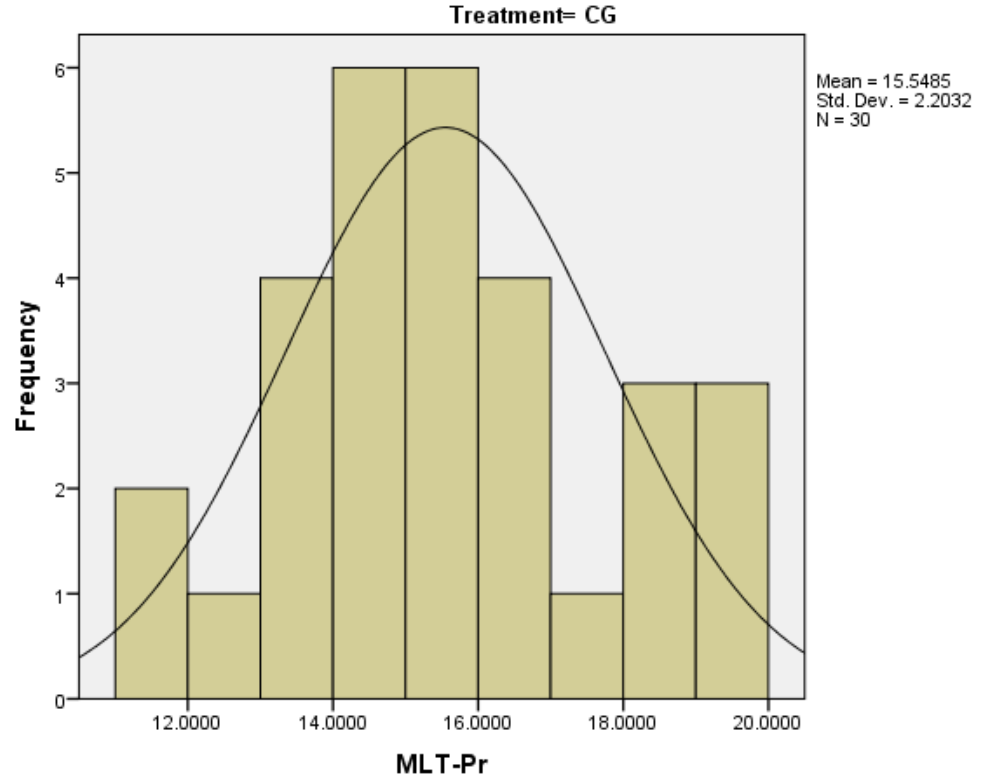


Figure 4.20. Pretest of MLT in the NF Group

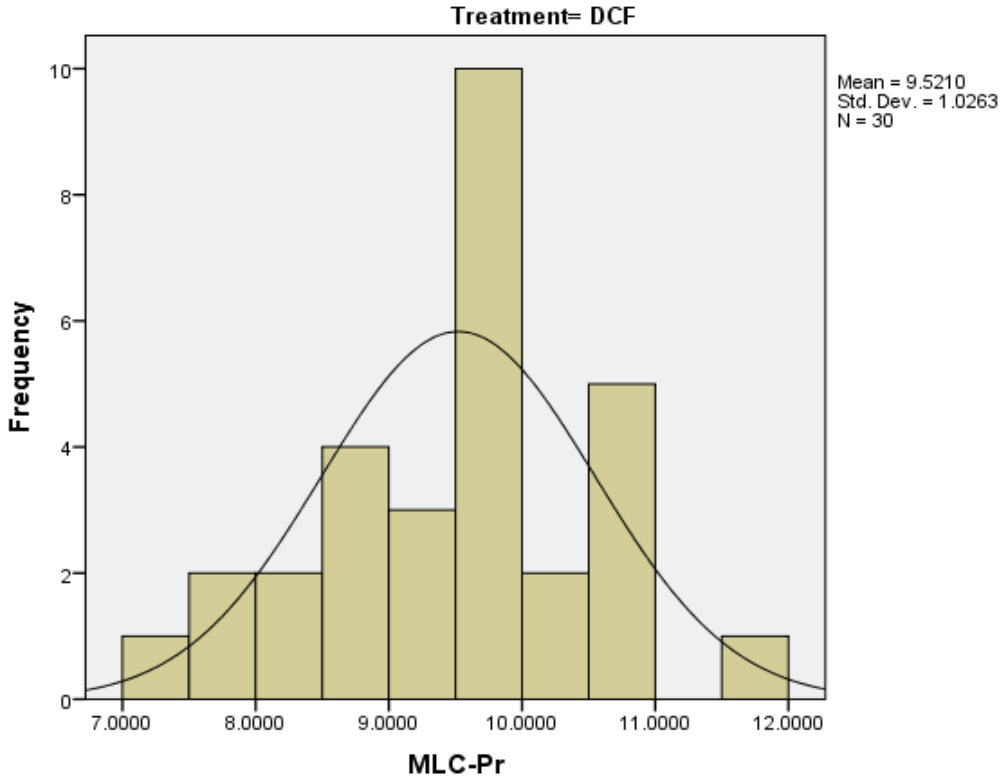


Figure 4.21. Pretest of MLC in the DCF Group

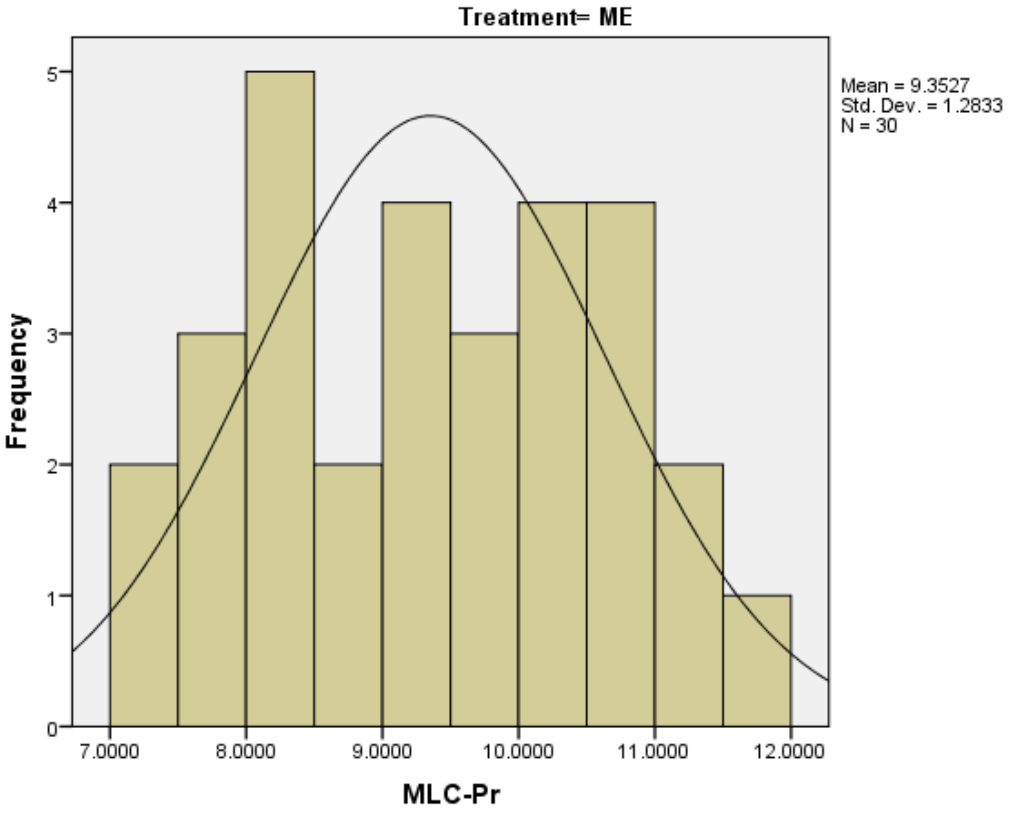


Figure 4.22. Pretest of MLC in the ME Group

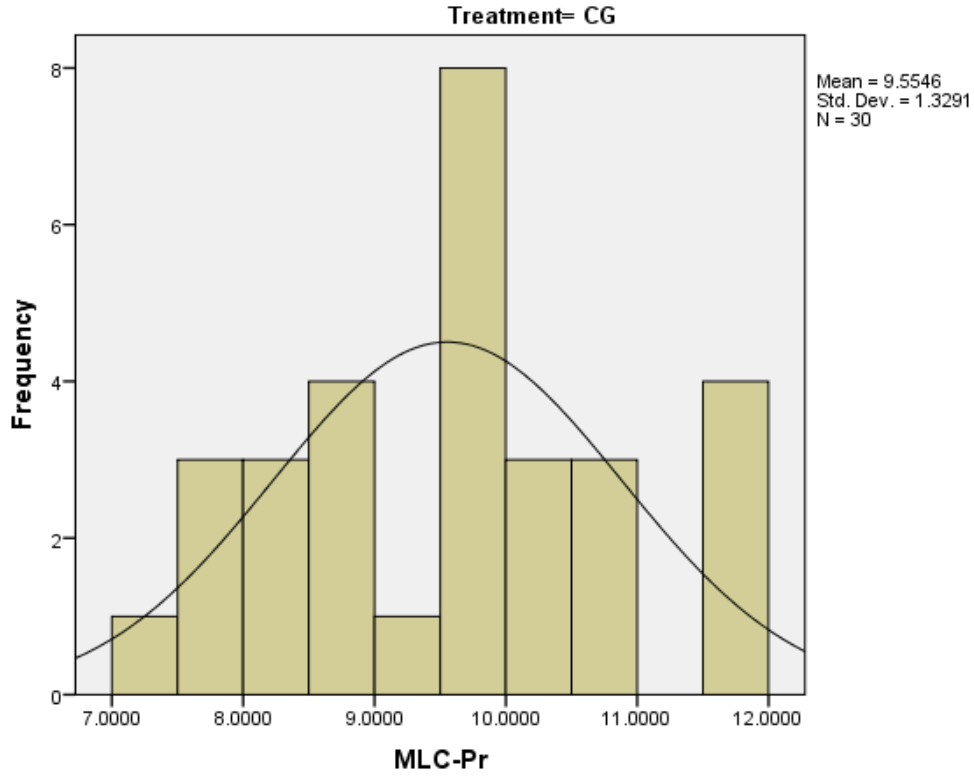


Figure 4.23. Pretest of MLC in the NF Group

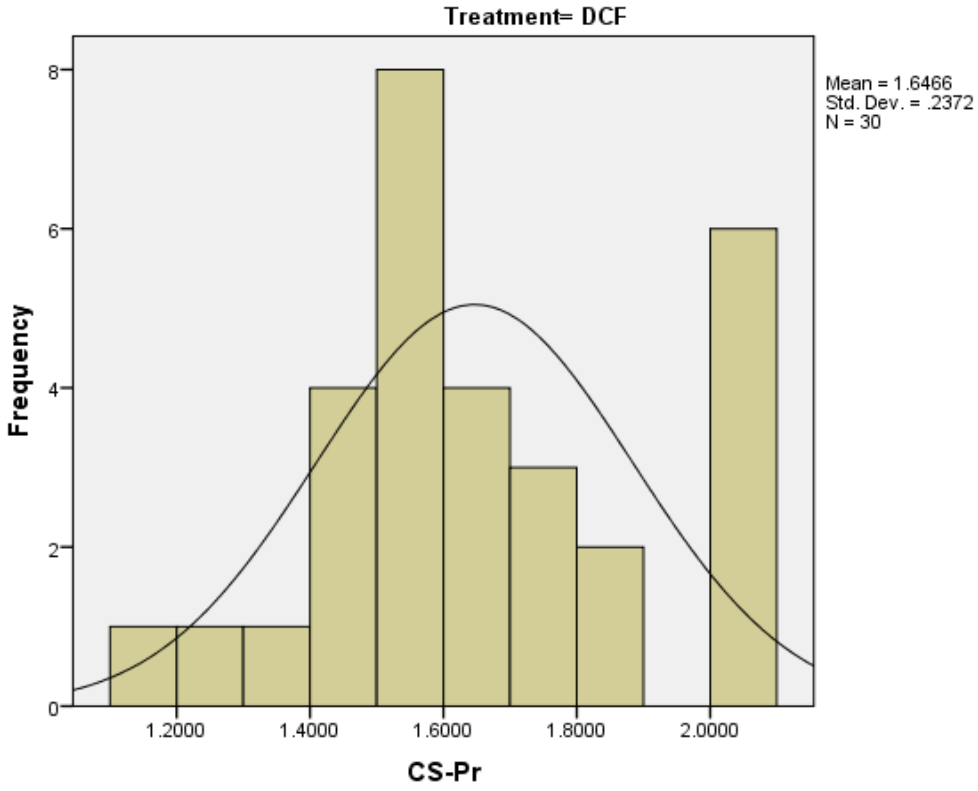


Figure 4.24. Pretest of C/S in the DCF Group

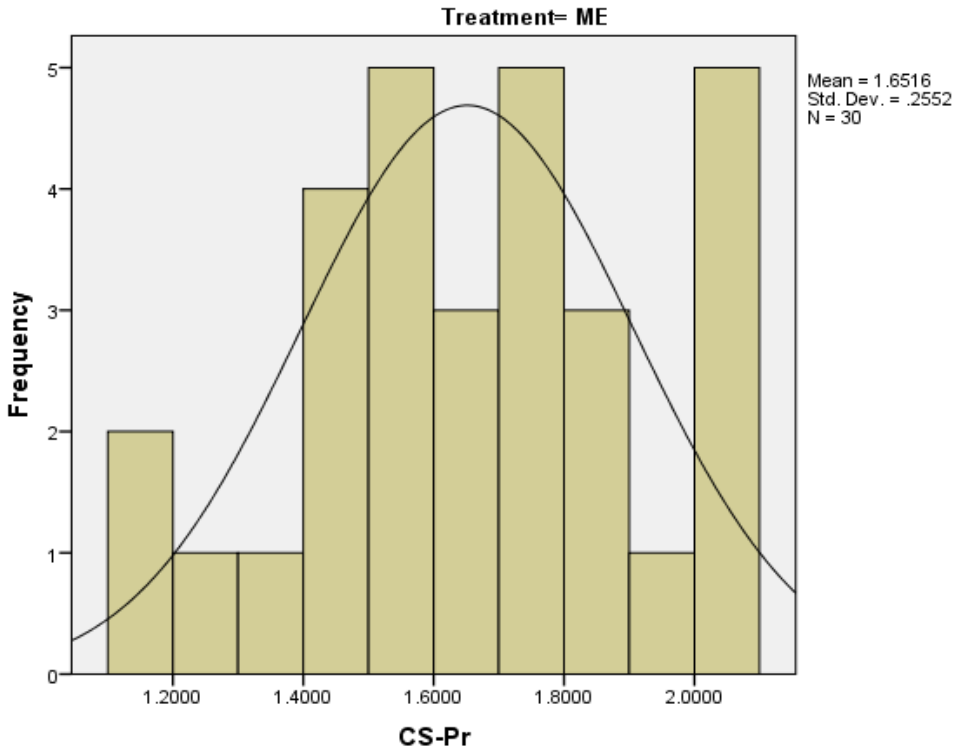


Figure 4.25. Pretest of C/S in the ME Group

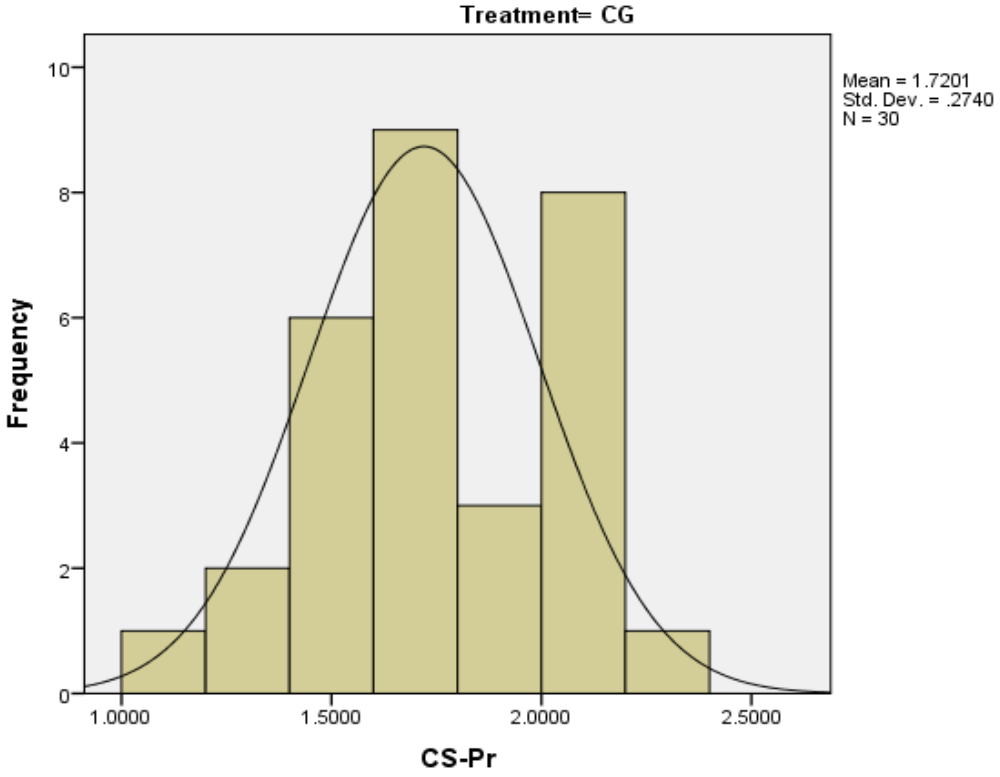


Figure 4.26. Pretest of C/S in the NF Group

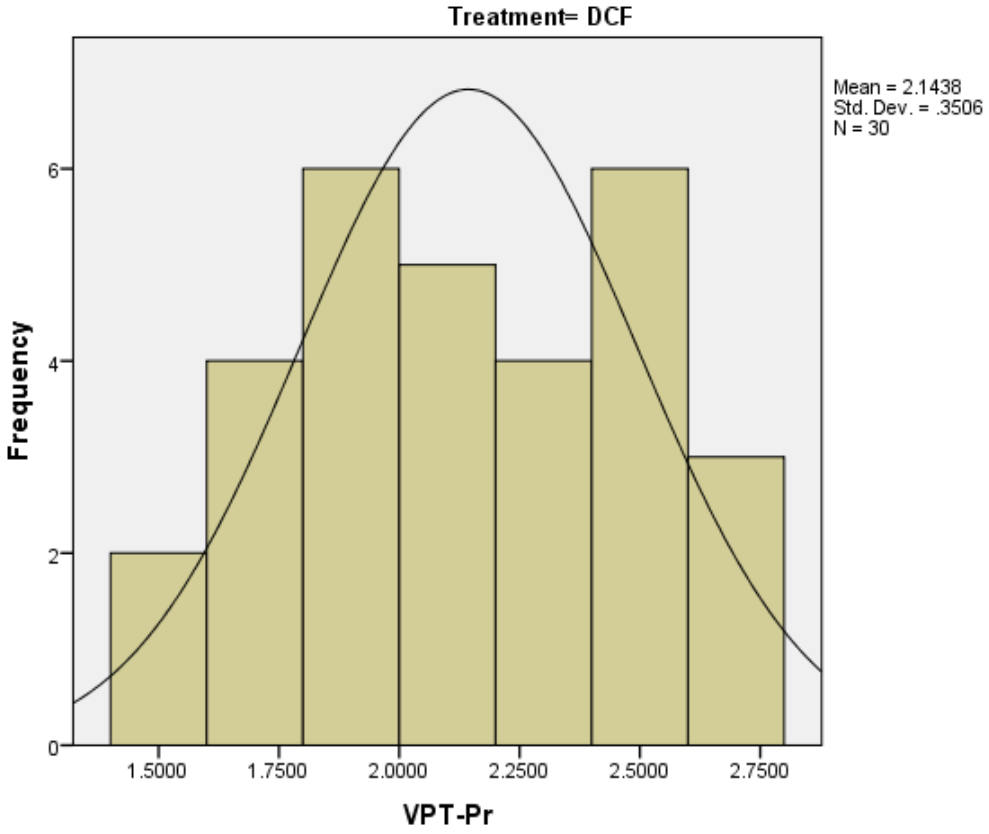


Figure 4.27. Pretest of VP/T in the DCF Group

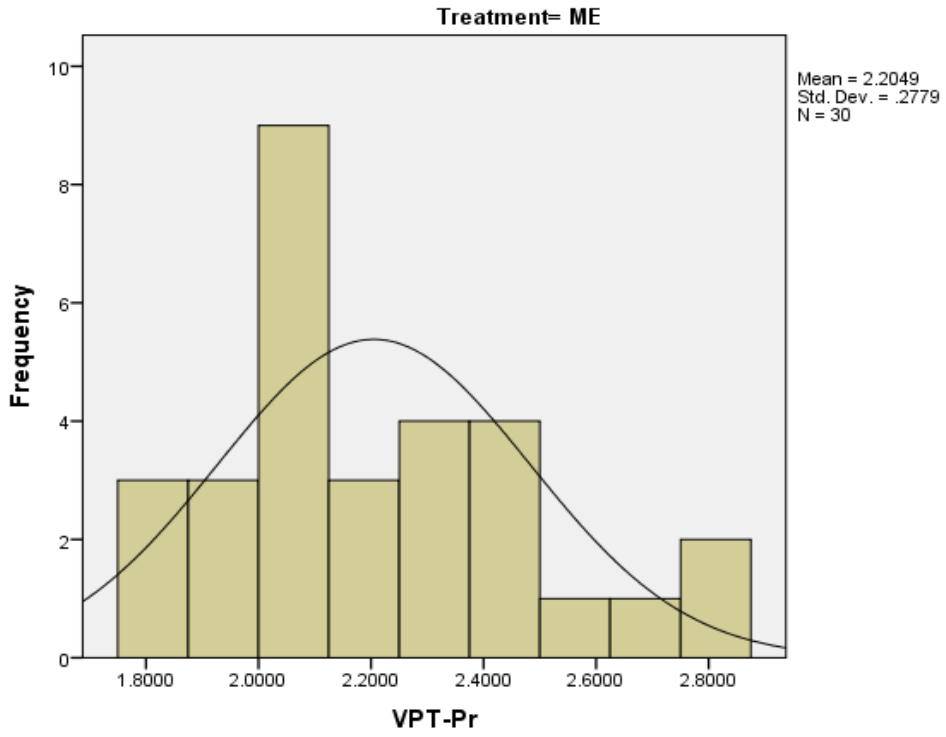


Figure 4.28. Pretest of VP/T in the ME Group

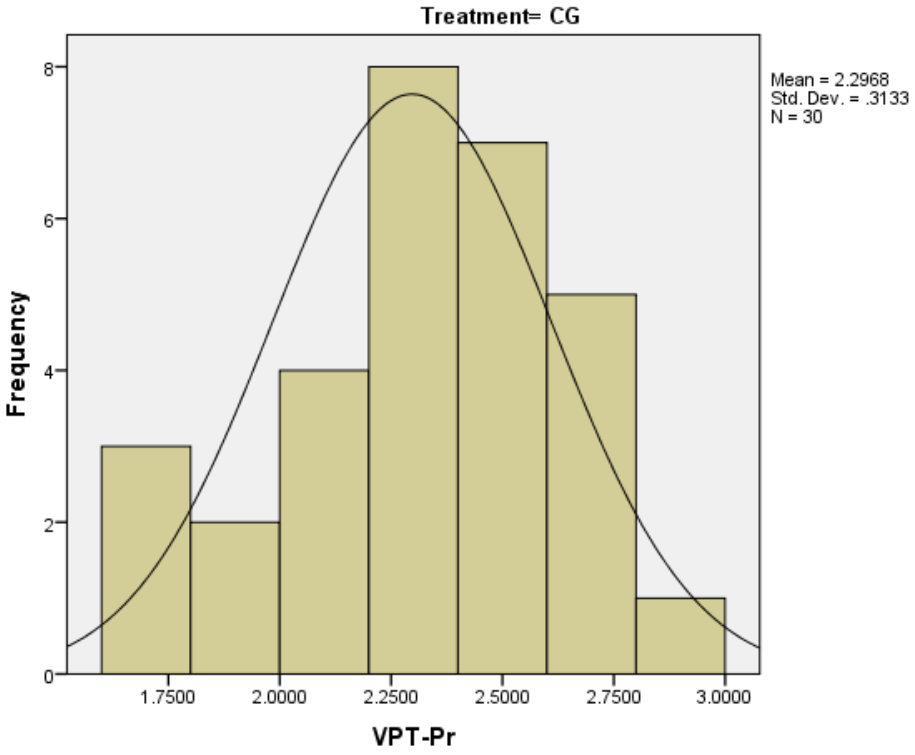


Figure 4.29. Pretest of VP/T in the NF Group

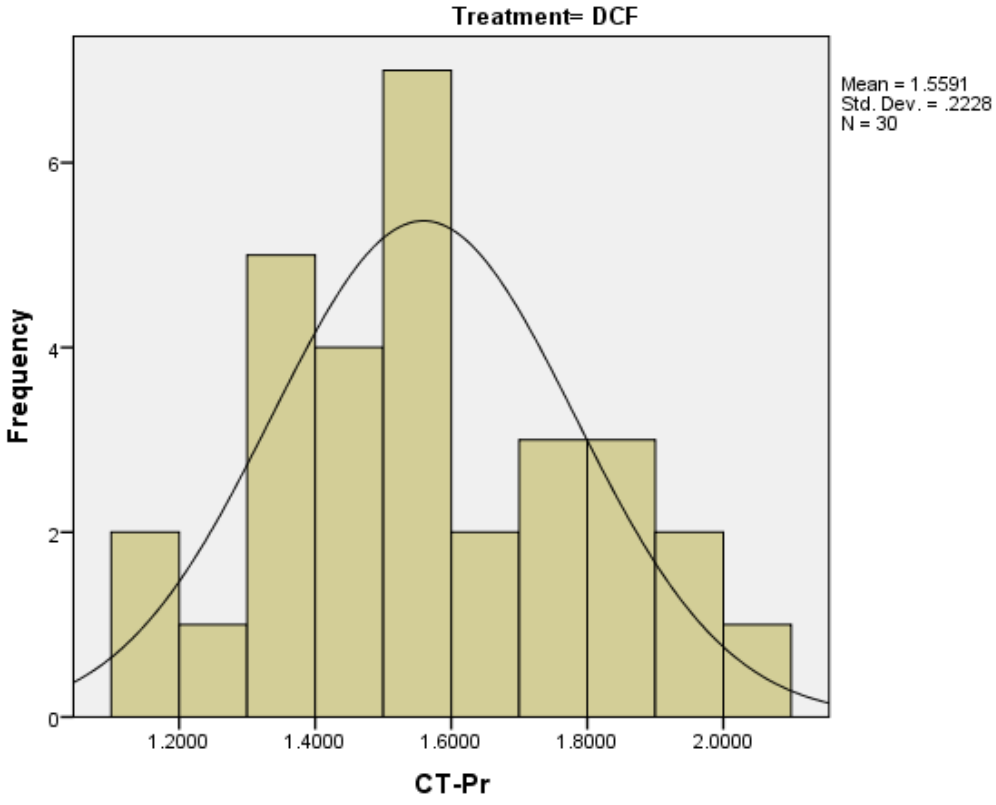


Figure 4.30. Pretest of C/T in the DCF Group

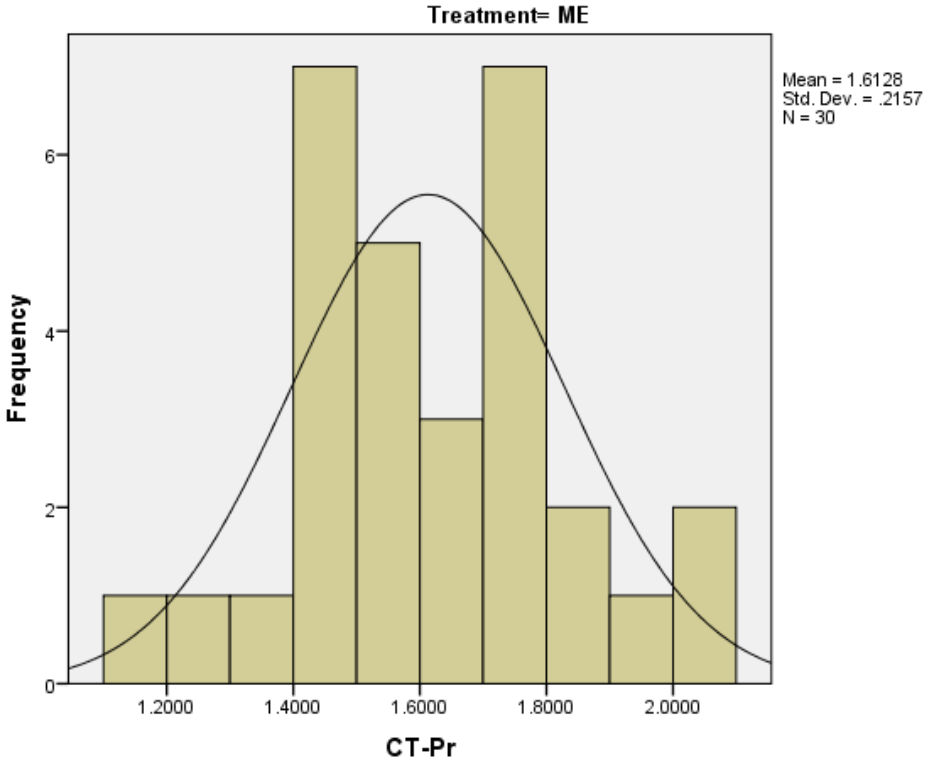


Figure 4.31. Pretest of C/T in the ME Group

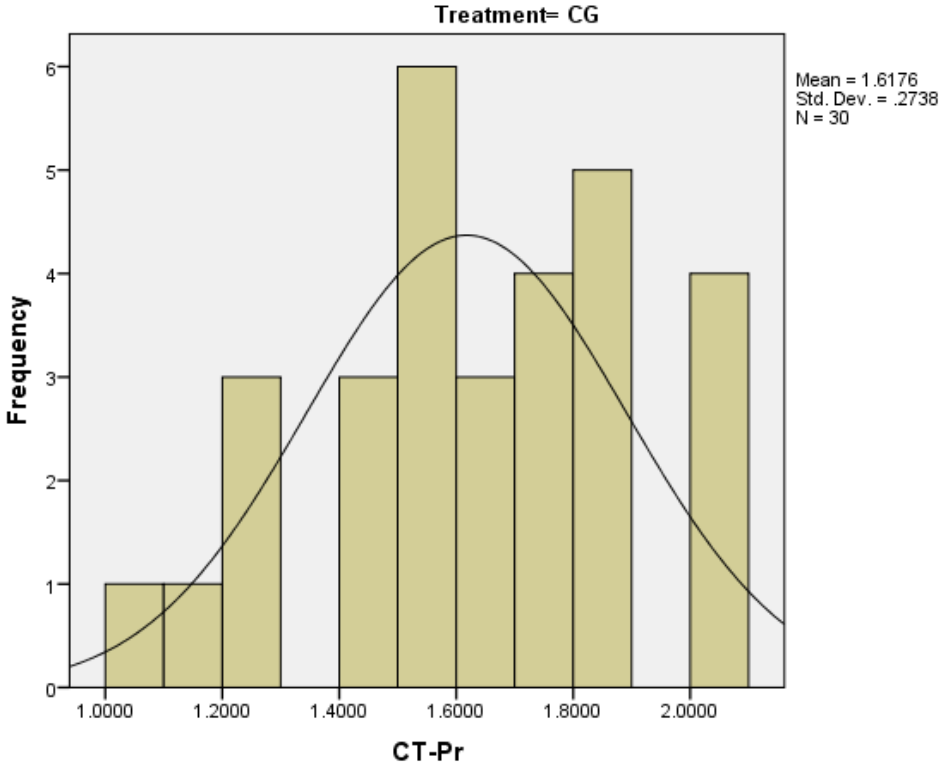


Figure 4.32. Pretest of C/T in the NF Group

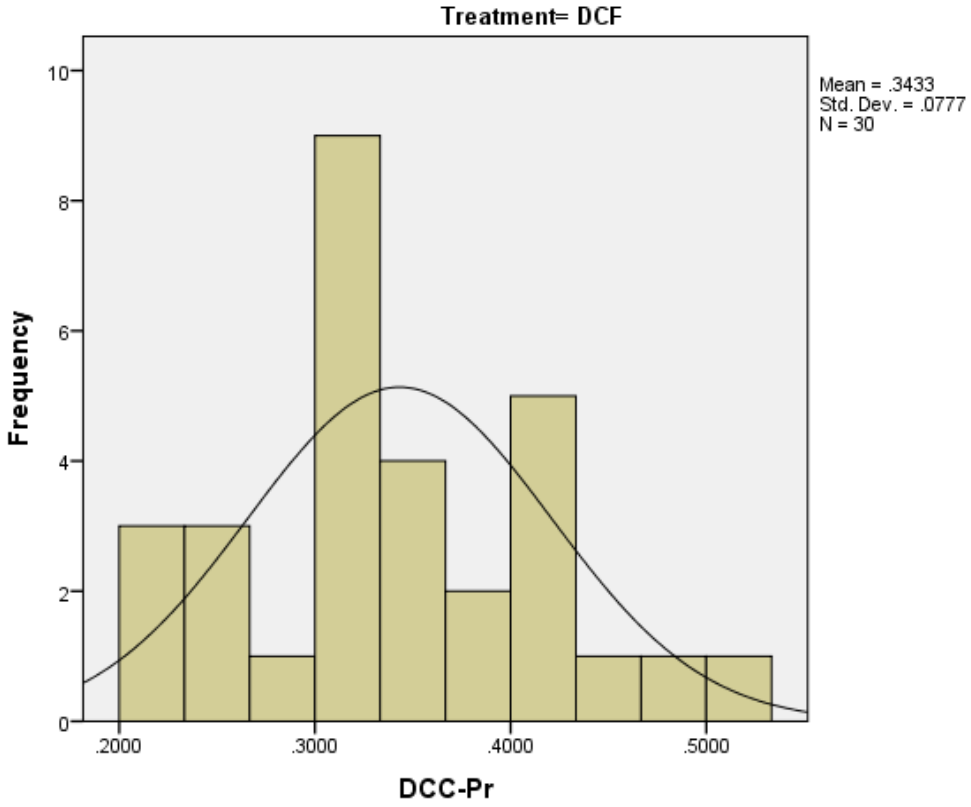


Figure 4.33. Pretest of DC/C in the DCF Group

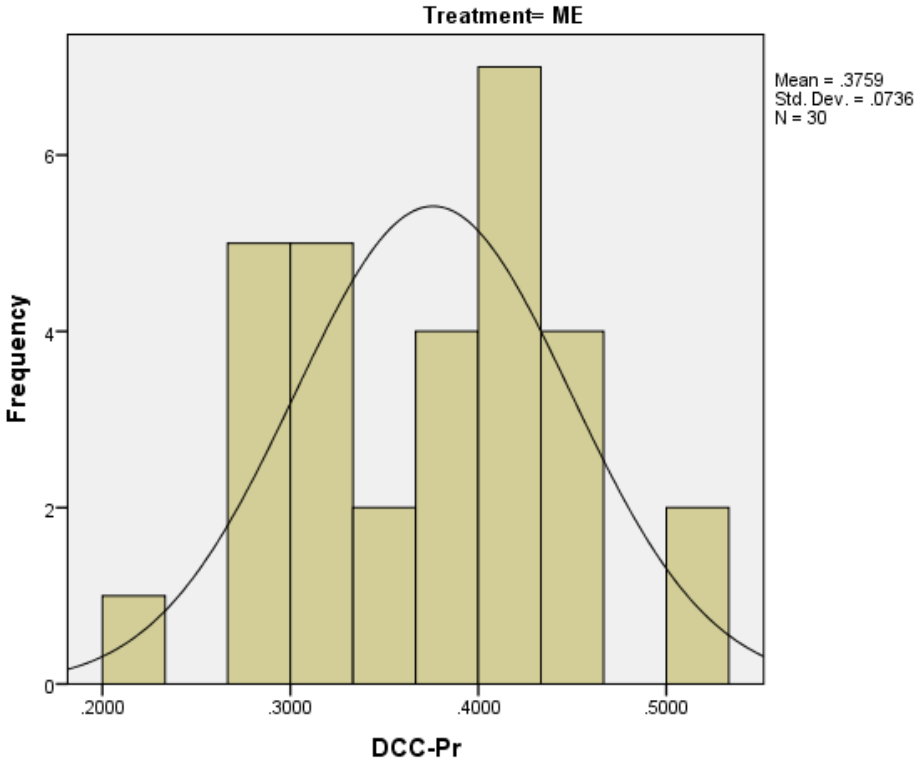


Figure 4.34. Pretest of DC/C in the ME Group

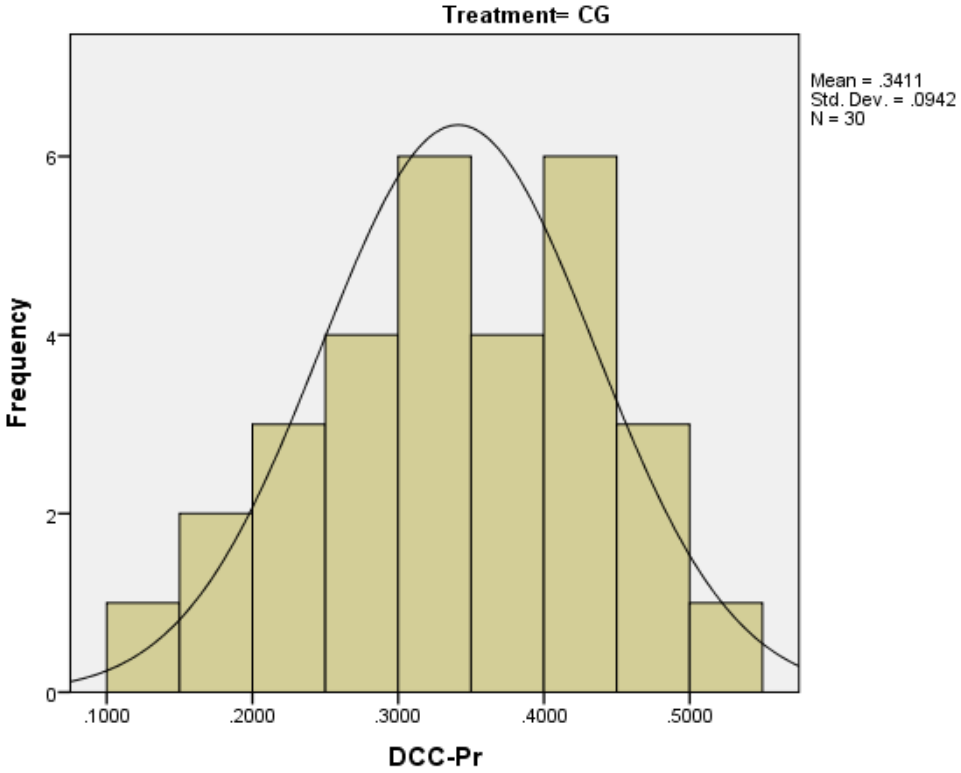


Figure 4.35. Pretest of DC/C in the NF Group

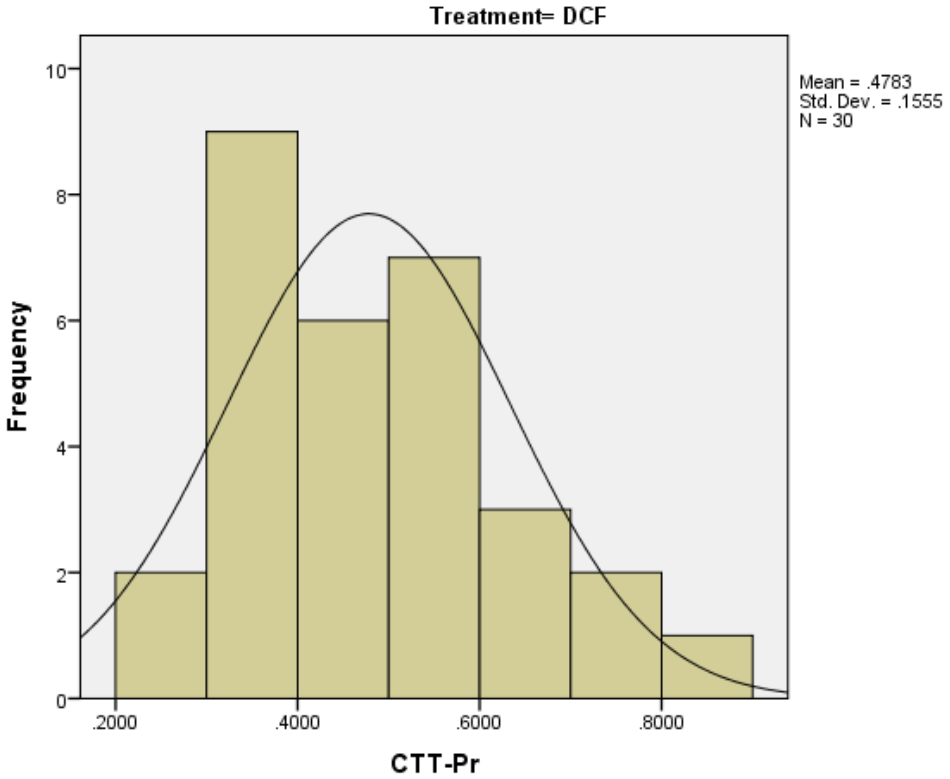


Figure 4.36. Pretest of CT/T in the DCF Group

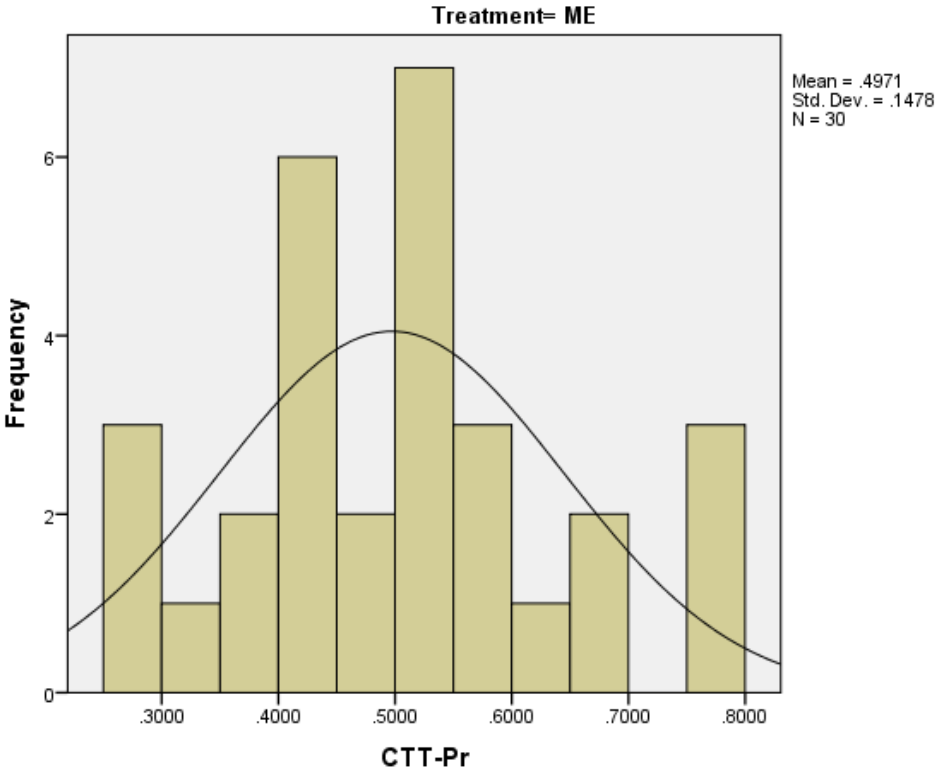


Figure 4.37. Pretest of CT/T in the ME Group

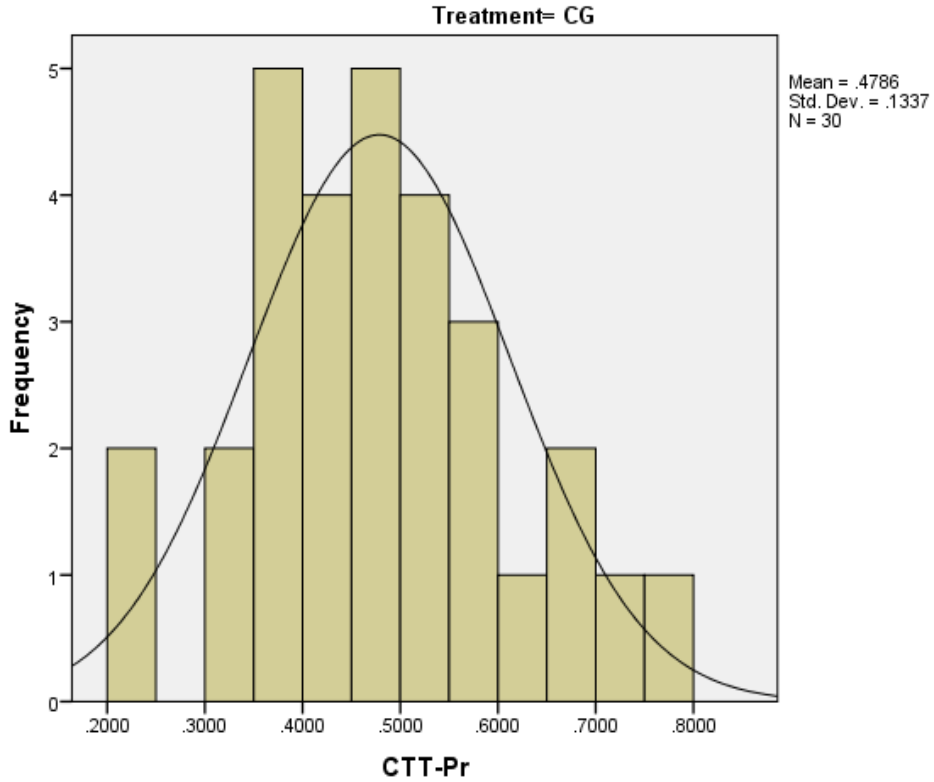


Figure 4.38. Pretest of CT/T in the NF Group

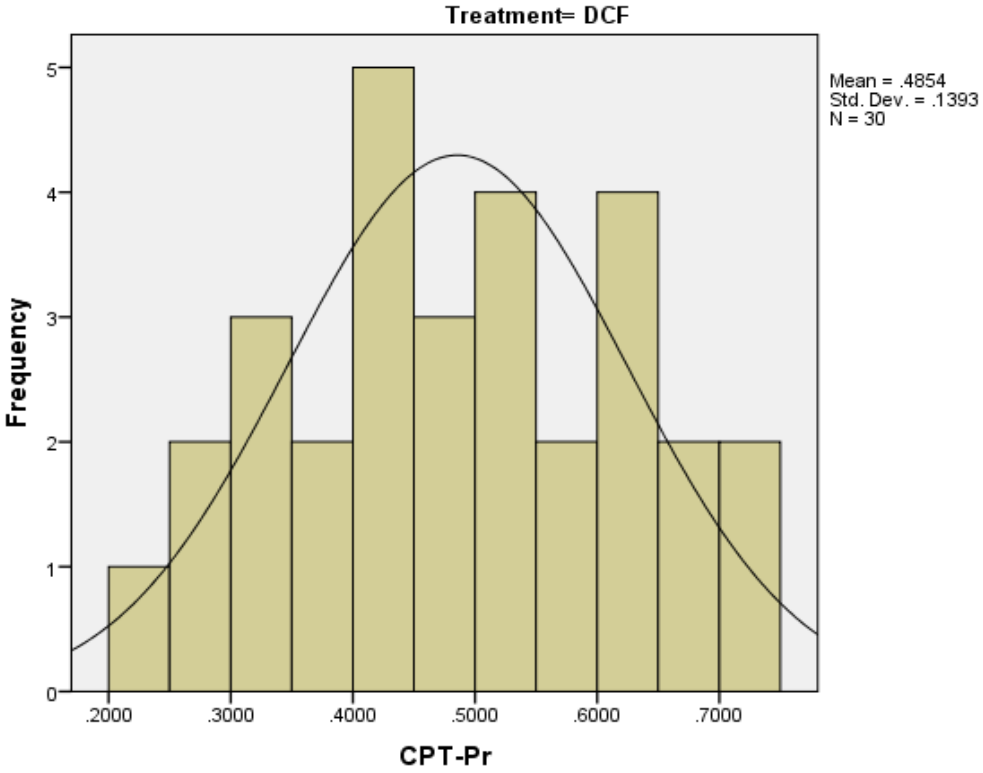


Figure 4.39. Pretest of CP/T in the DCF Group

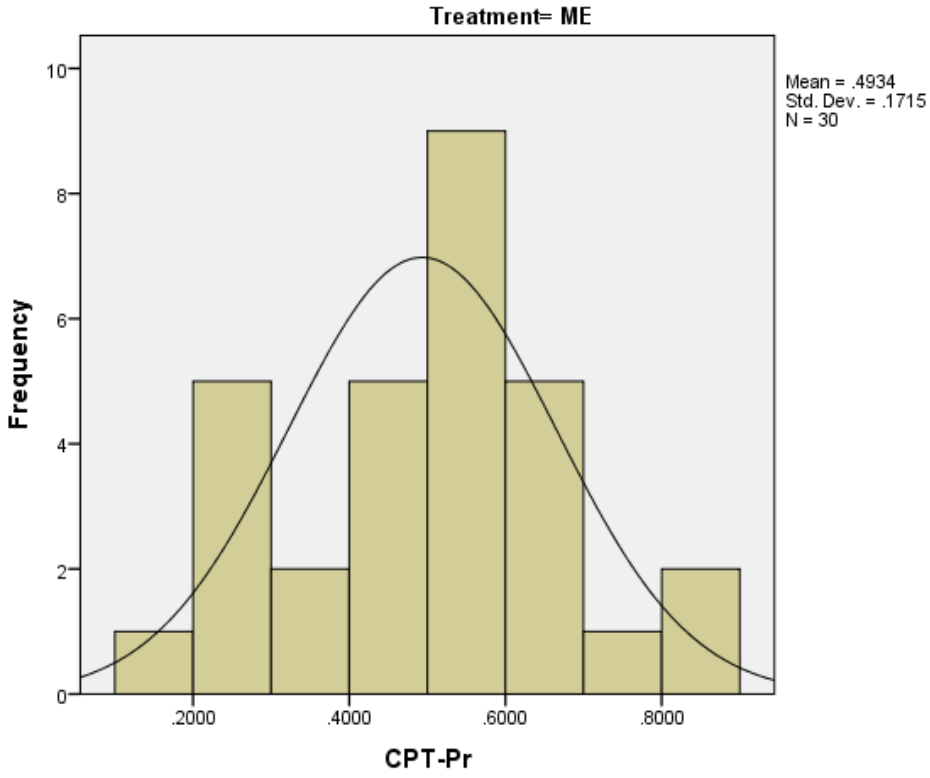


Figure 4.40. Pretest of CP/T in the ME Group

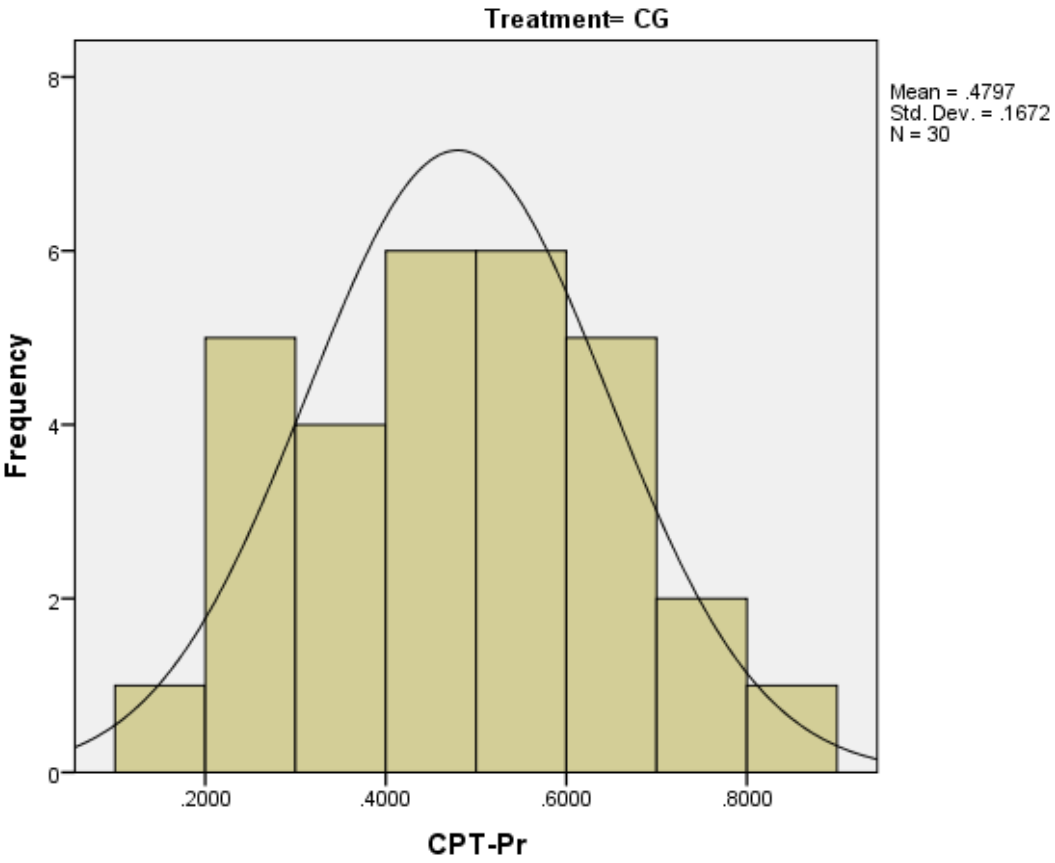


Figure 4.41. Pretest of CP/T in the NF Group

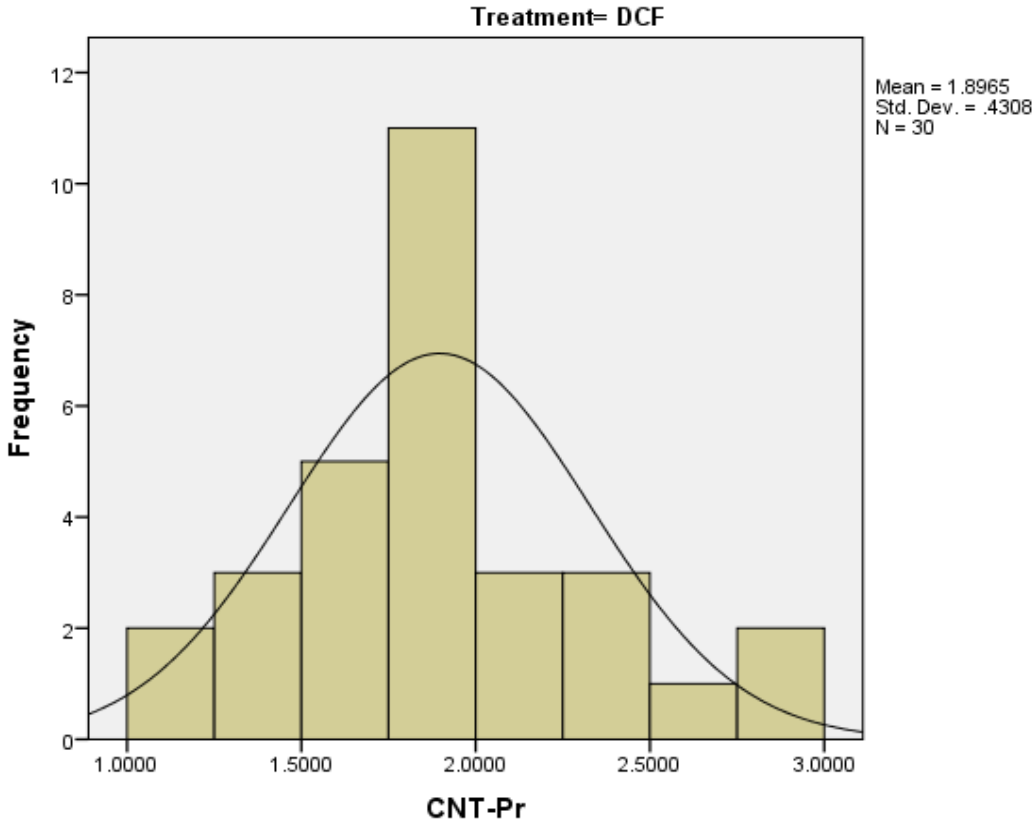


Figure 4.42. Pretest of CN/T in the DCF Group

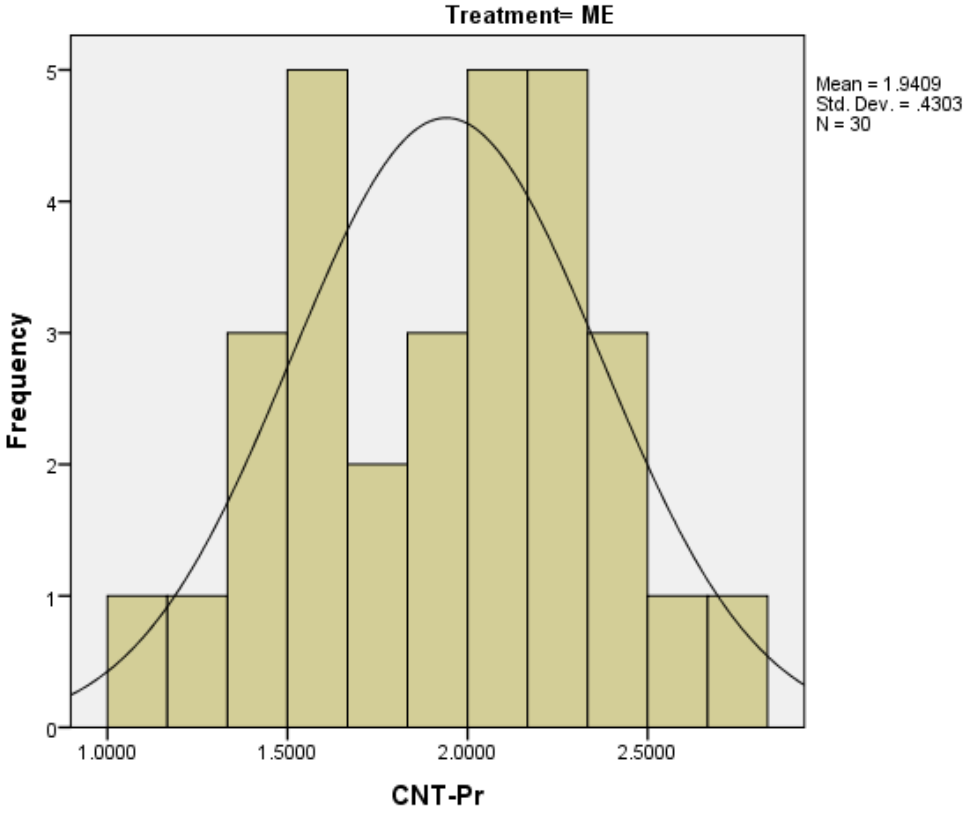


Figure 4.43. Pretest of CN/T in the ME Group

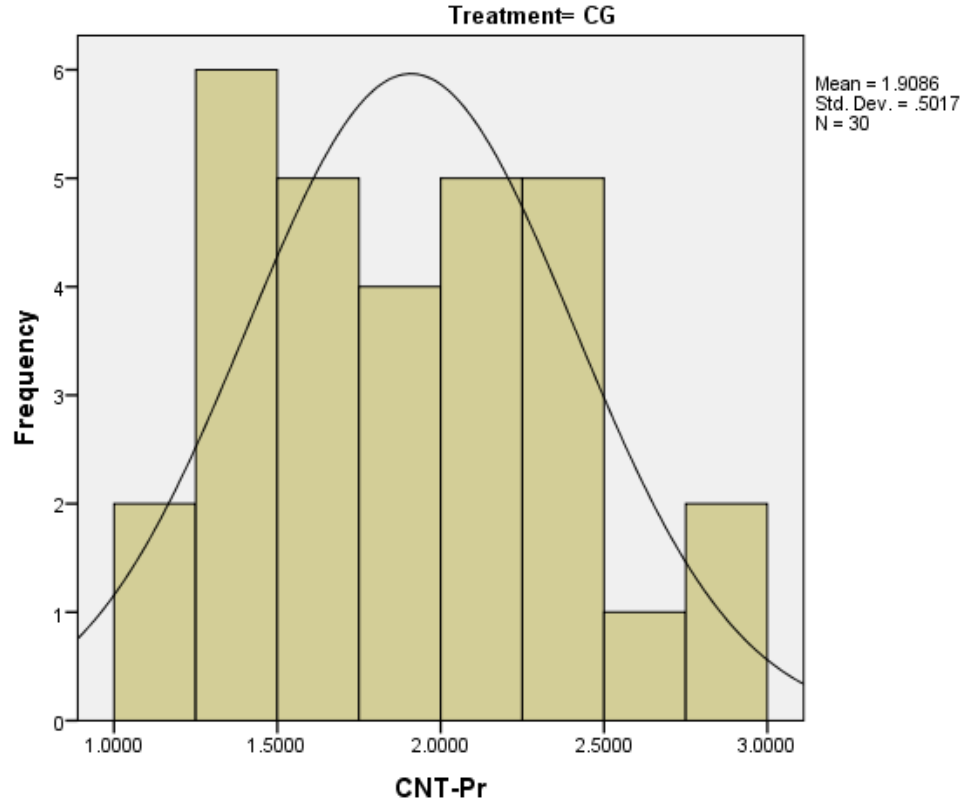


Figure 4.44. Pretest of CN/T in the NF Group

Appendix 7. Posttest of Syntactic Complexity

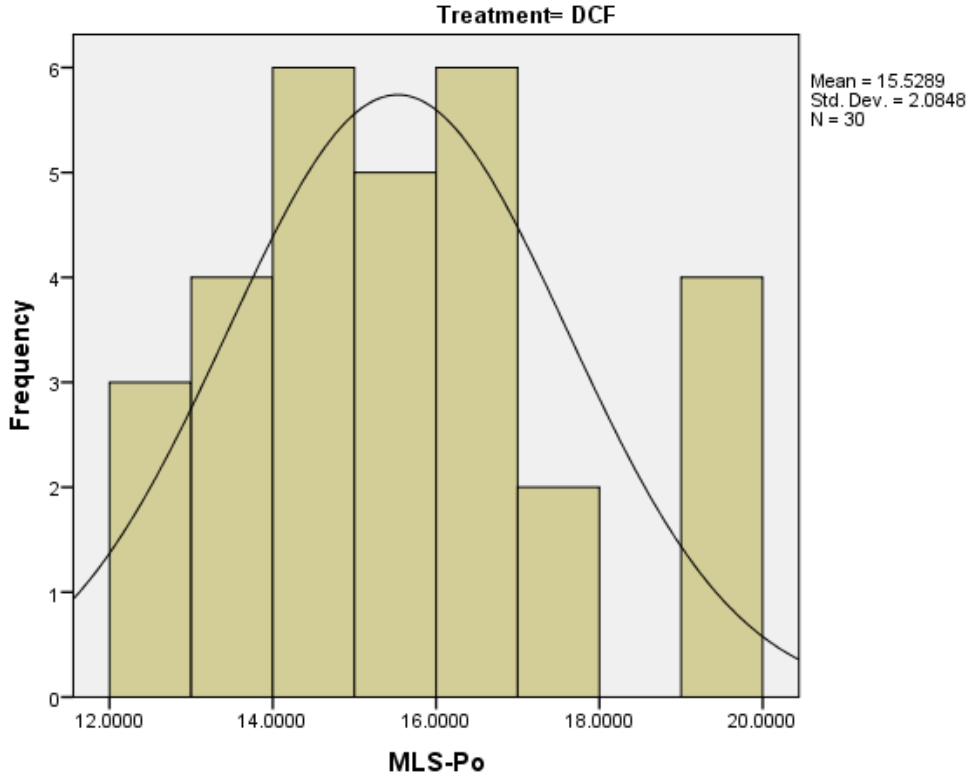


Figure 4.45. Posttest of MLS in the DCF Group

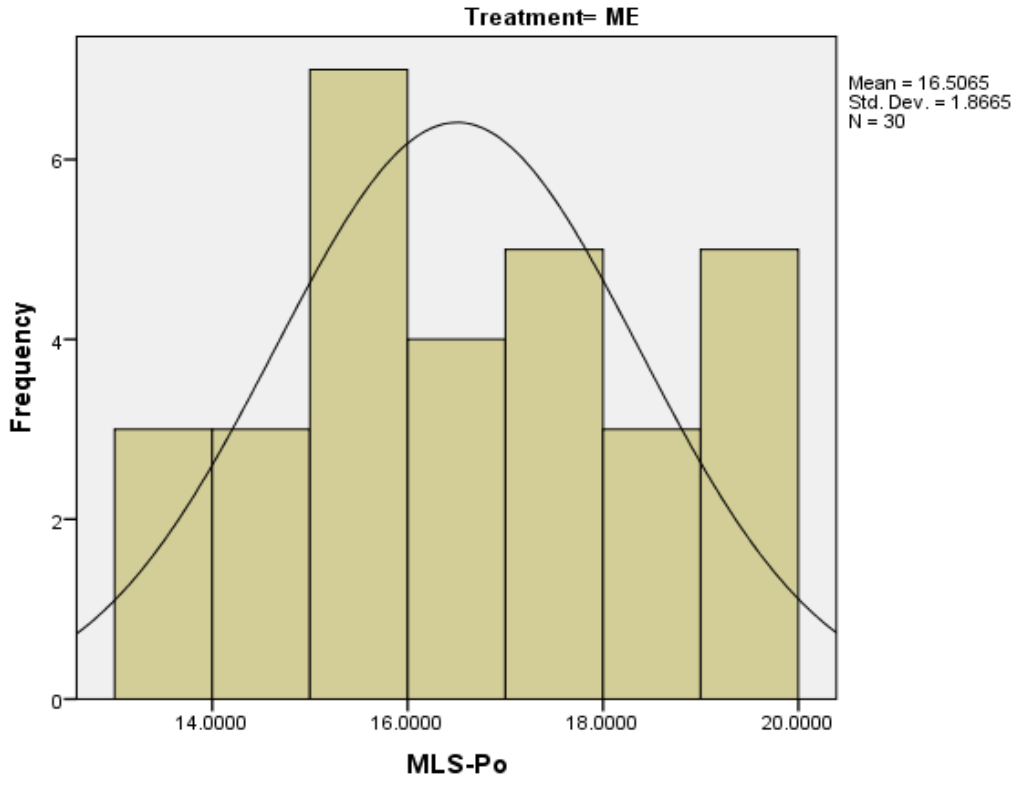


Figure 4.46. Posttest of MLS in the ME Group

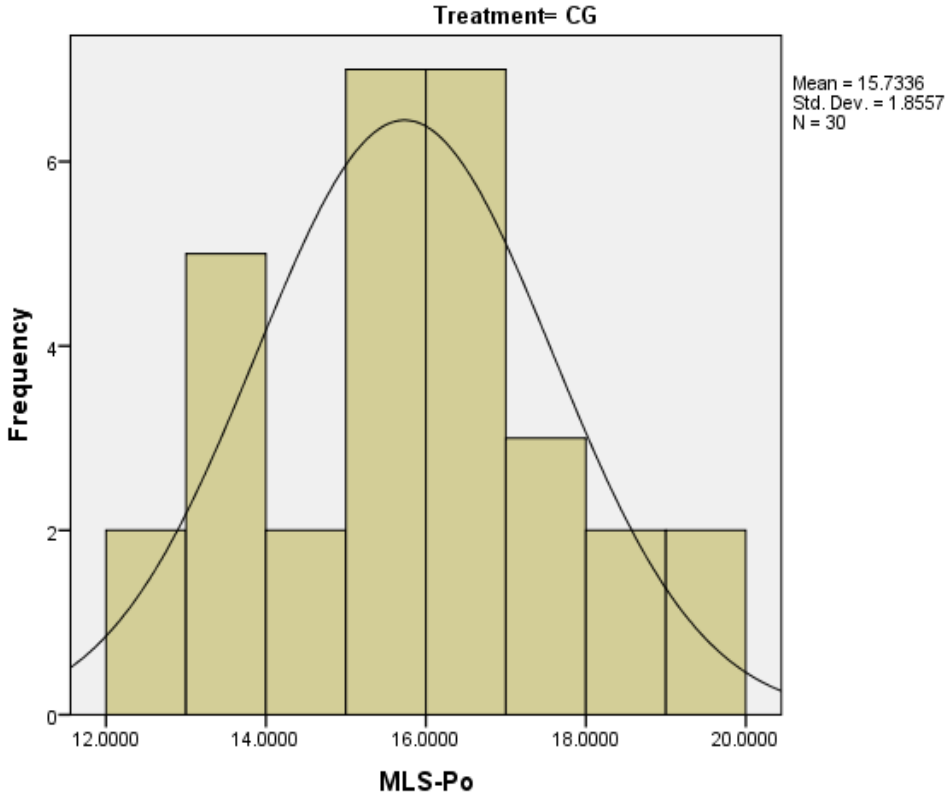


Figure 4.47. Posttest of MLS in the NF Group

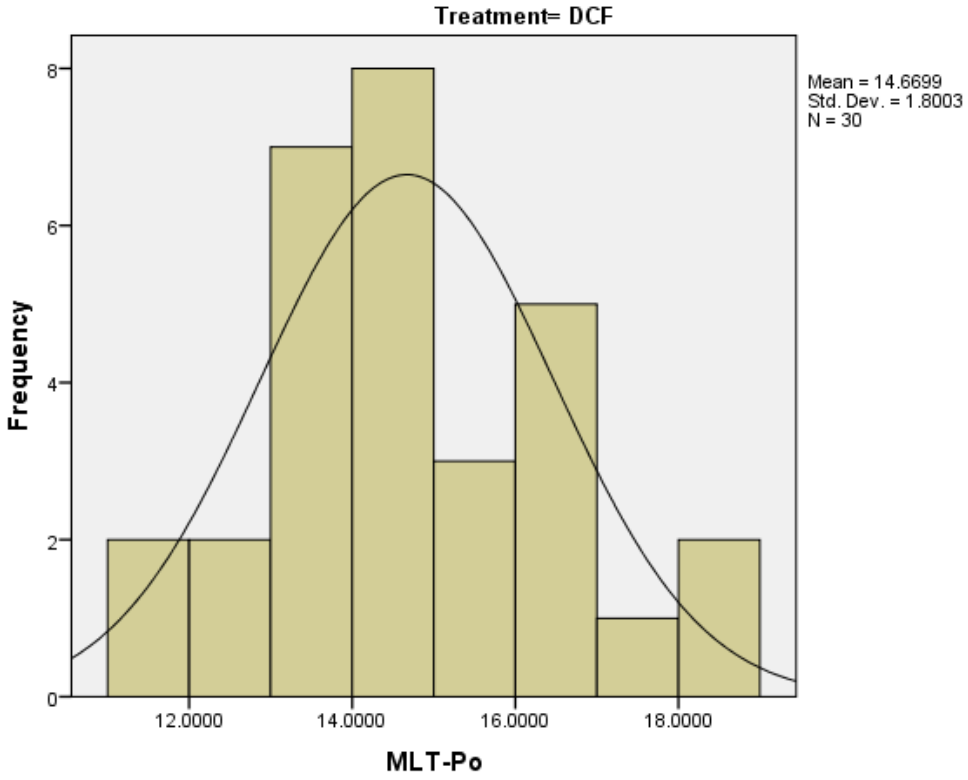


Figure 4.48. Posttest of MLT in the DCF Group

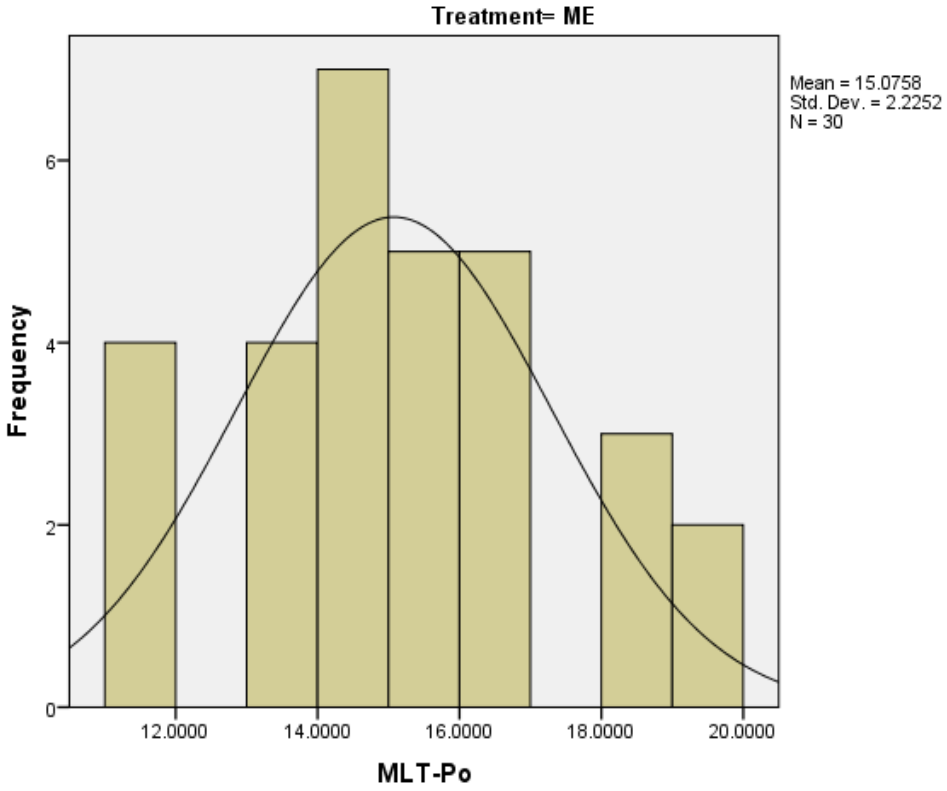


Figure 4.49. Posttest of MLT in the ME Group

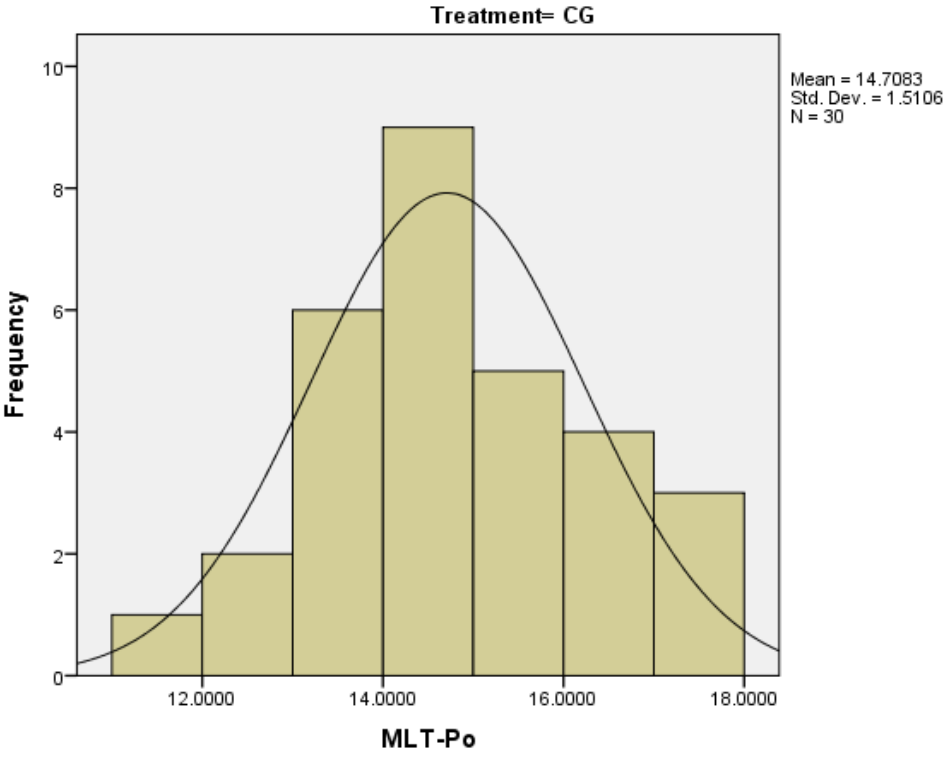


Figure 4.50. Posttest of MLT in the NF Group

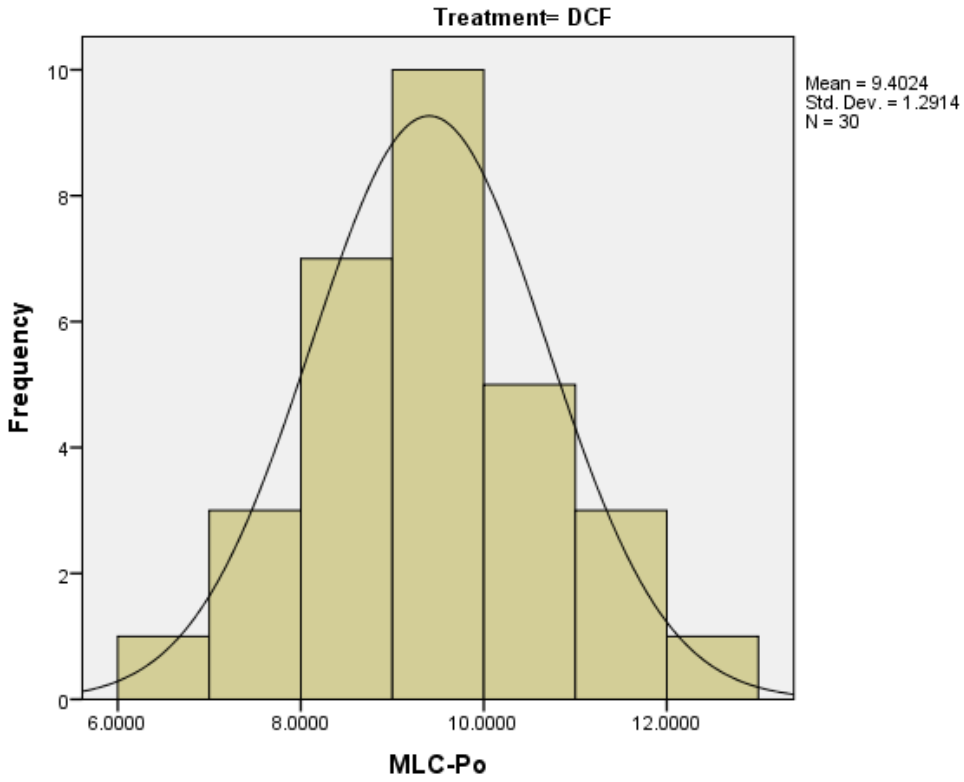


Figure 4.51. Posttest of MLC in the DCF Group

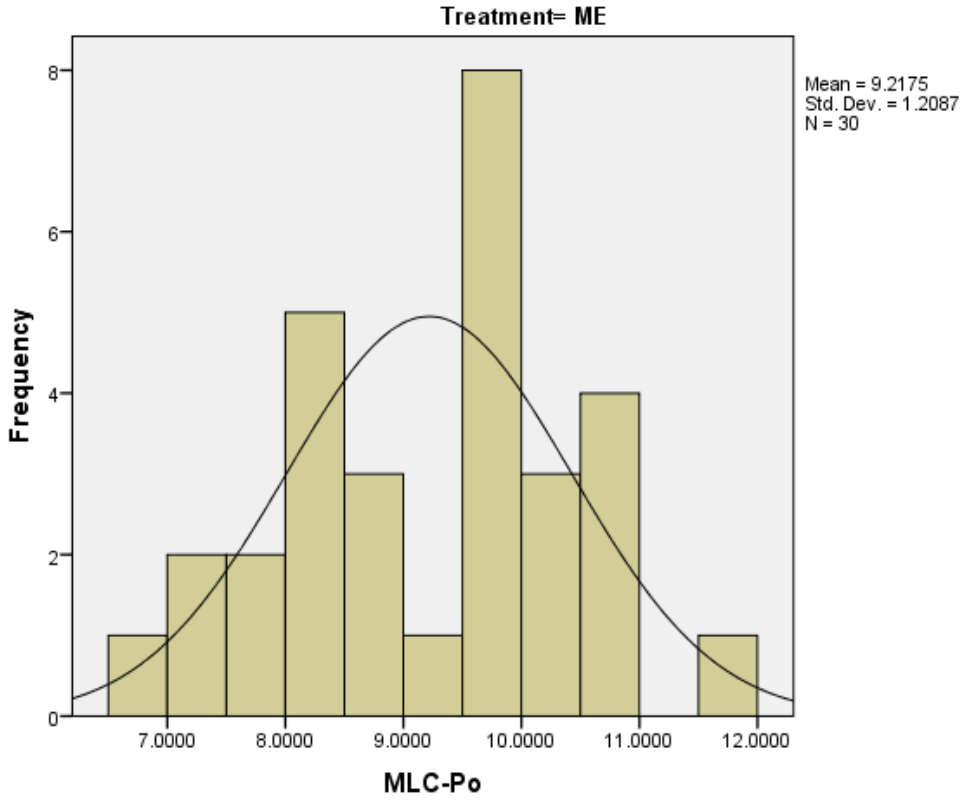


Figure 4.52. Posttest of MLC in the ME Group

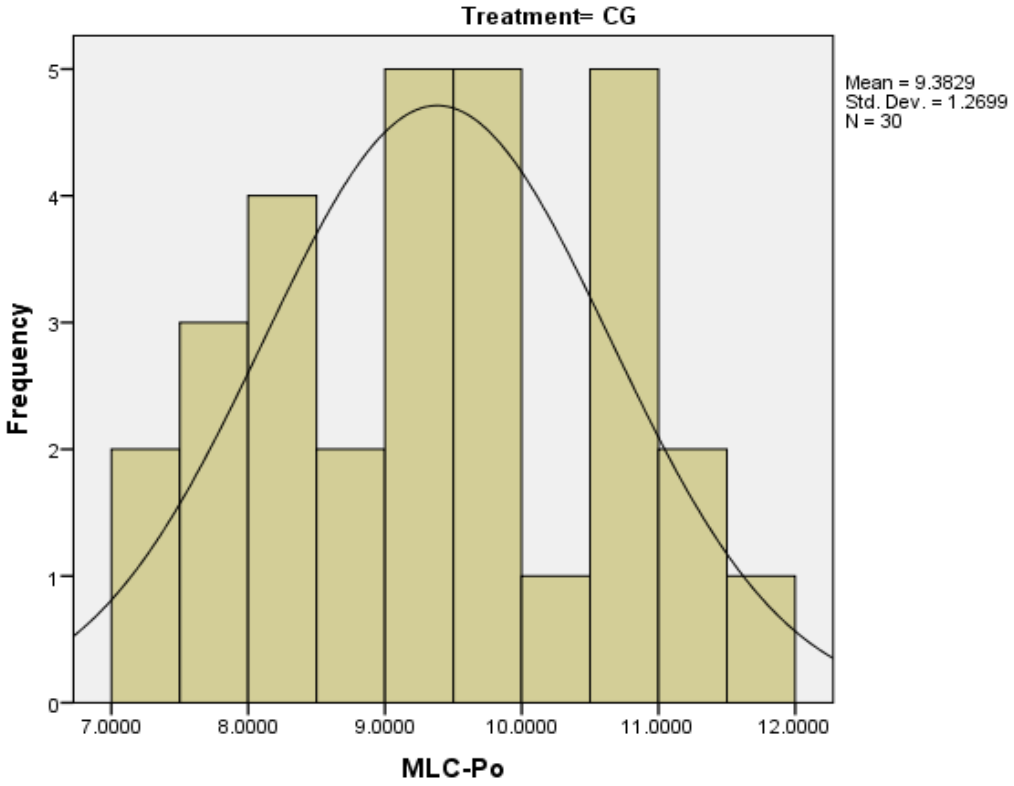


Figure 4.53. Posttest of MLC in the NF Group

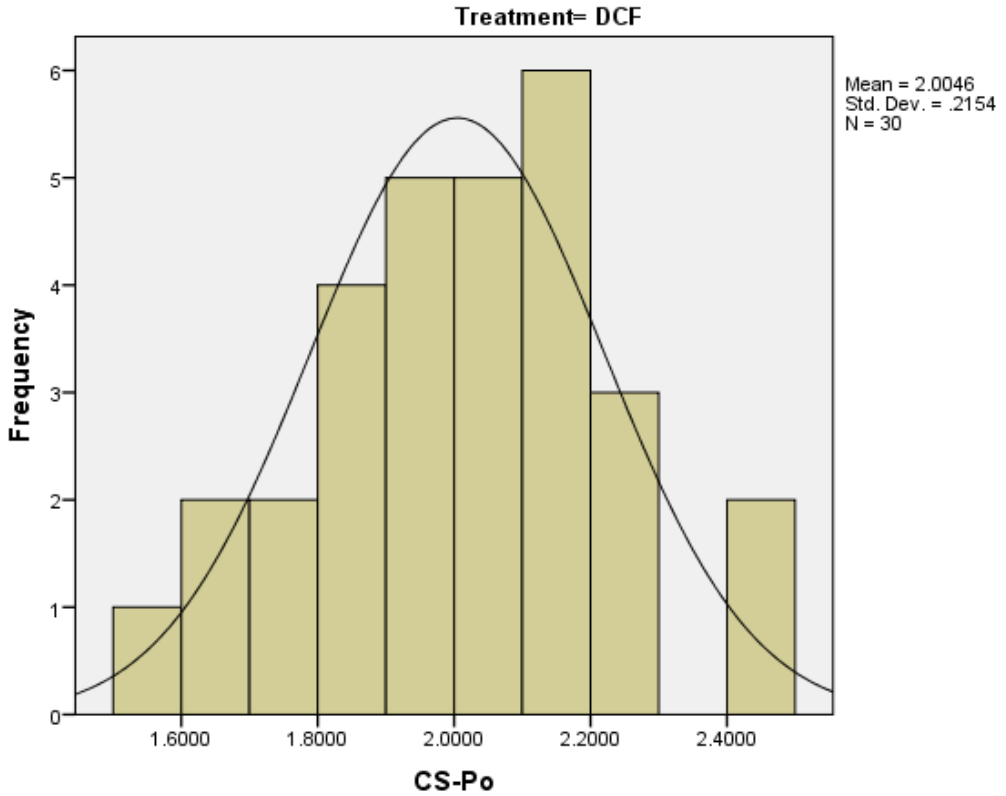


Figure 4.54. Posttest of C/S in the DCF Group

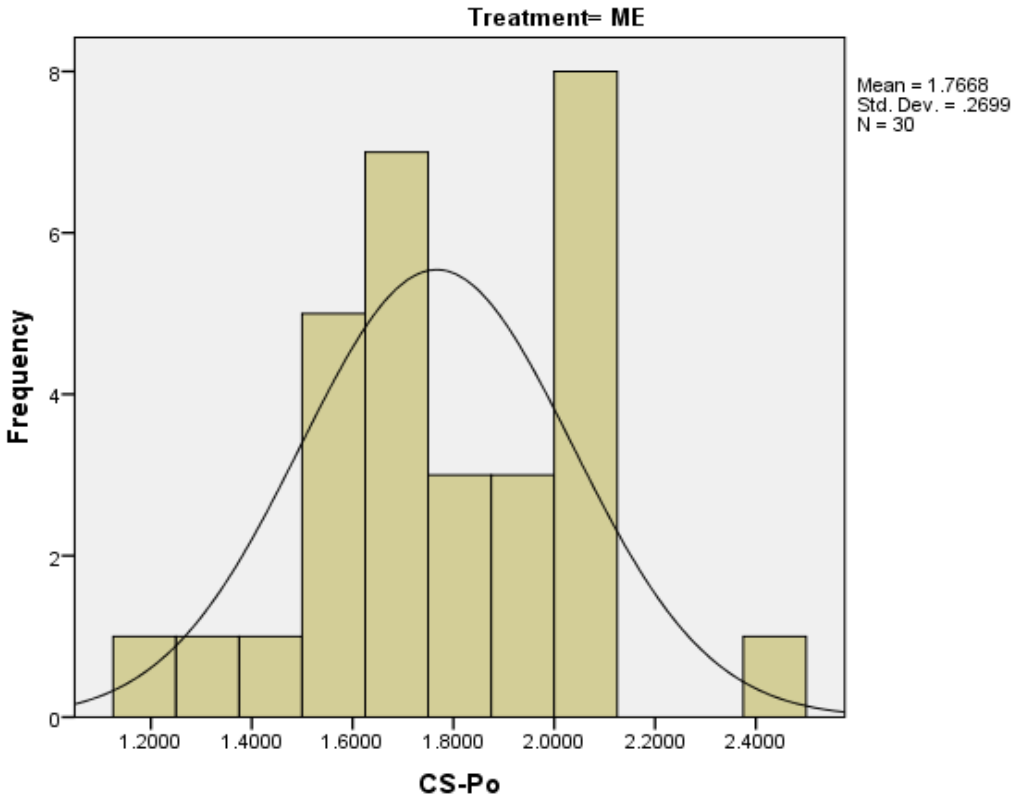


Figure 4.55. Posttest of C/S in the ME Group

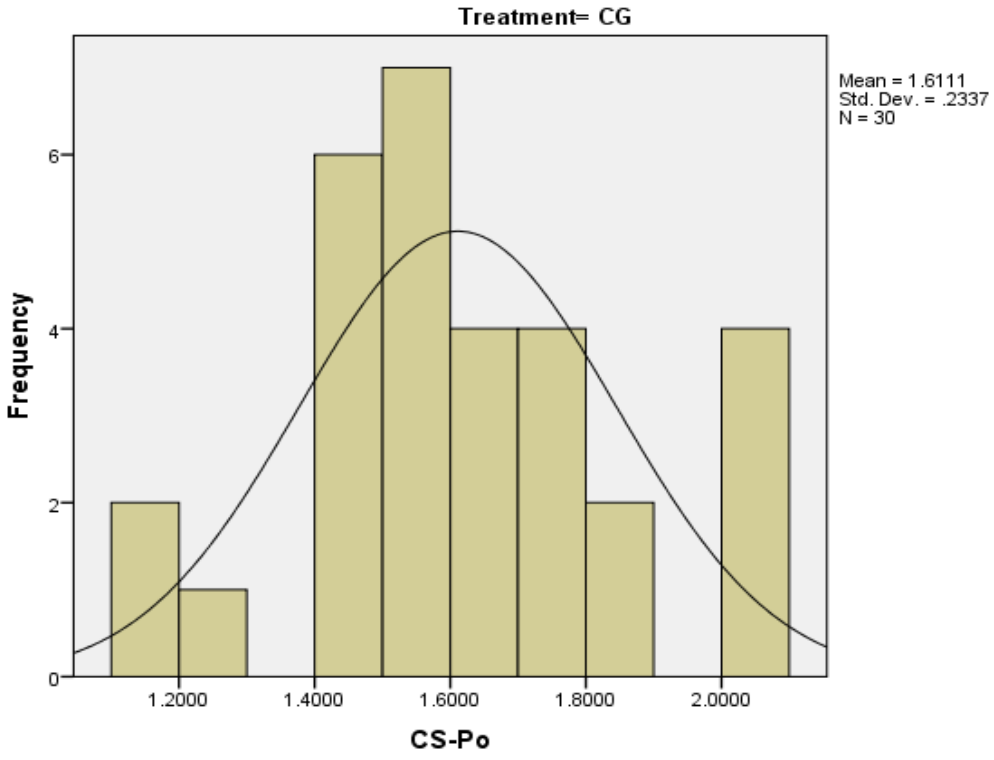


Figure 4.56. Posttest of C/S in the NF Group

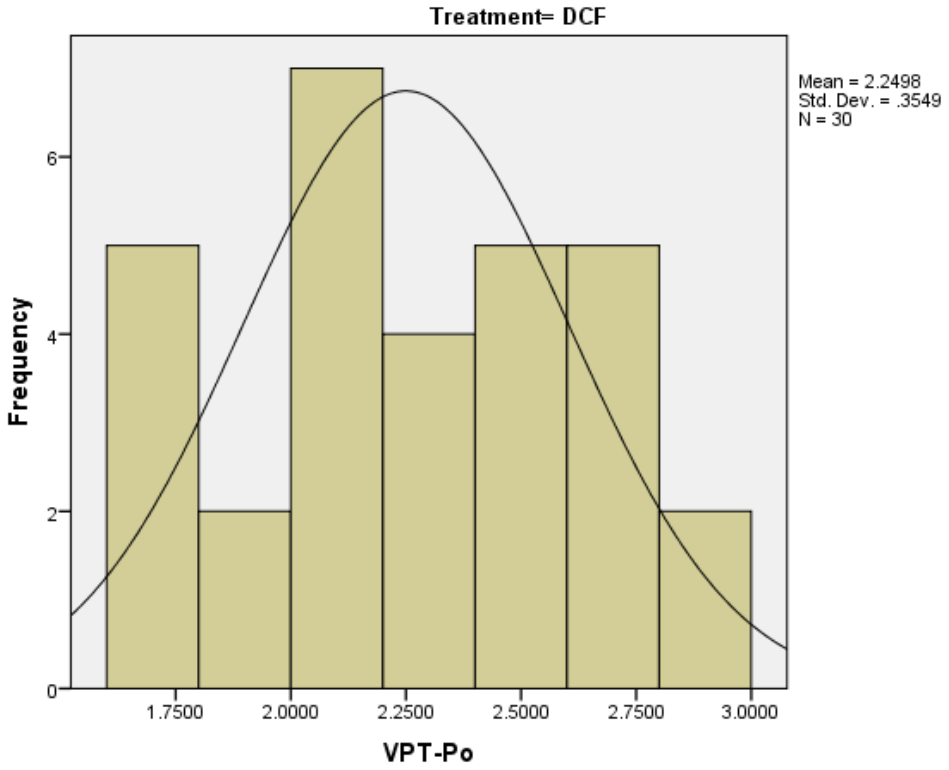


Figure 4.57. Posttest of VP/T in the DCF Group

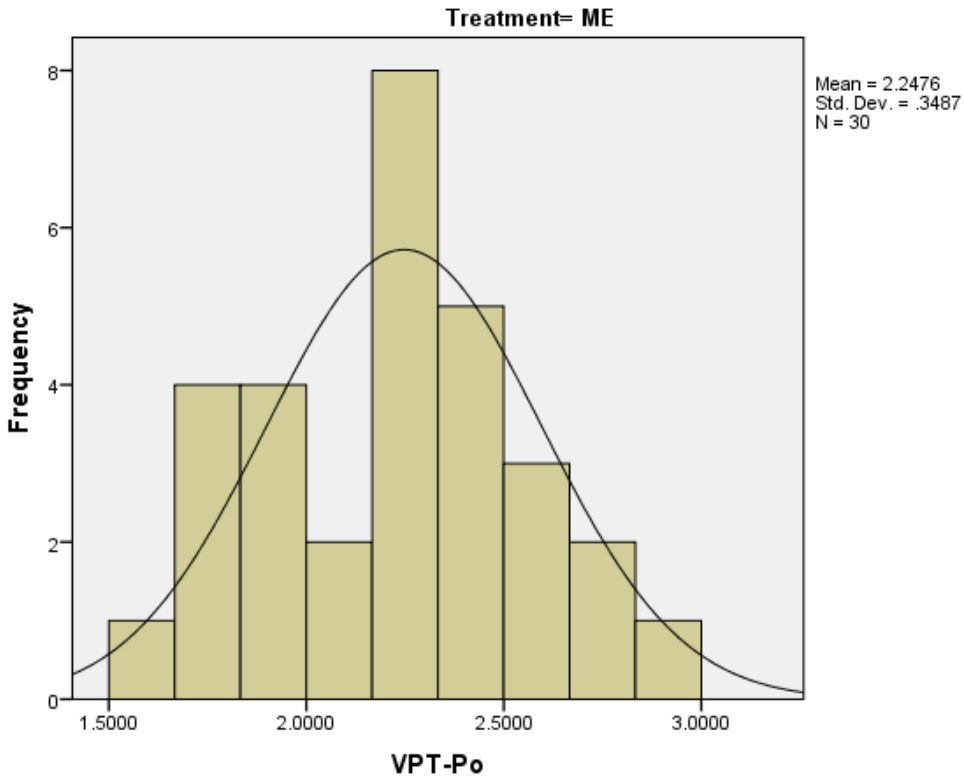


Figure 4.58. Posttest of VP/T in the ME Group

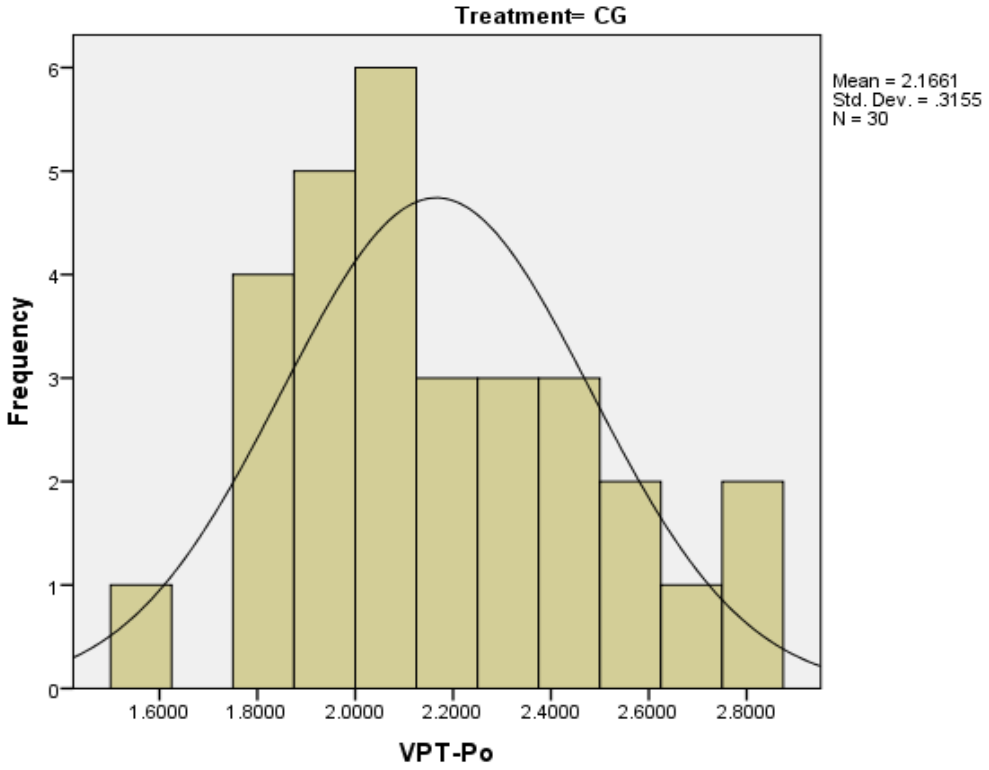


Figure 4.59. Posttest of VP/T in the NF Group

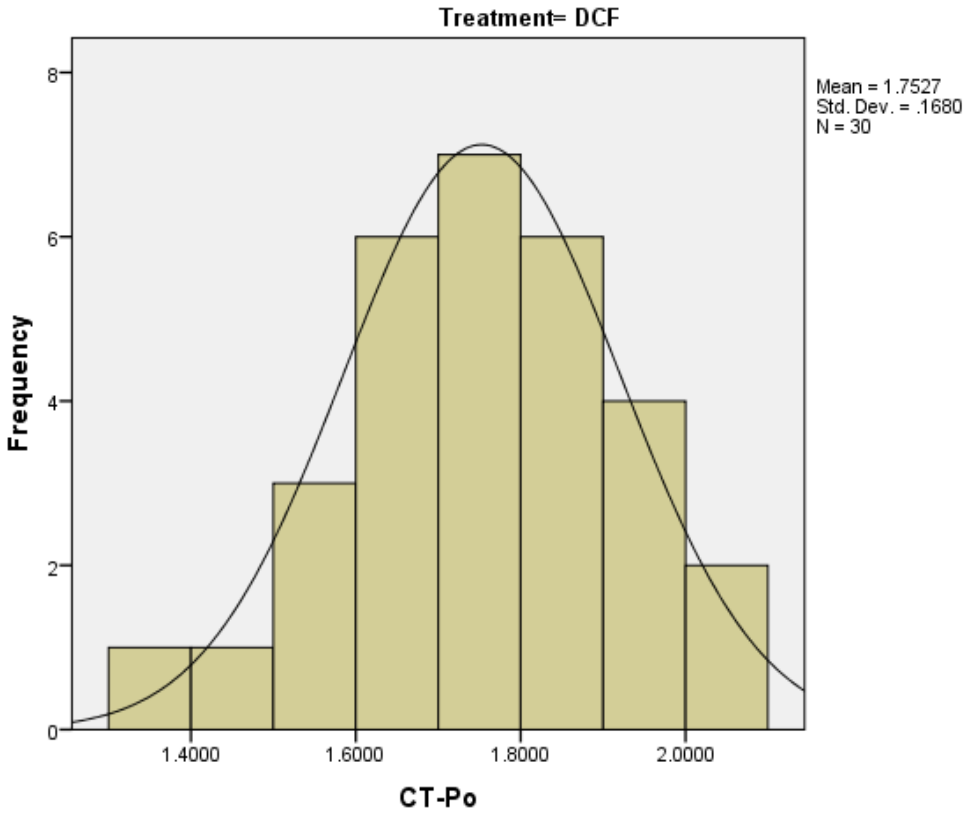


Figure 4.60. Posttest of C/T in the DCF Group

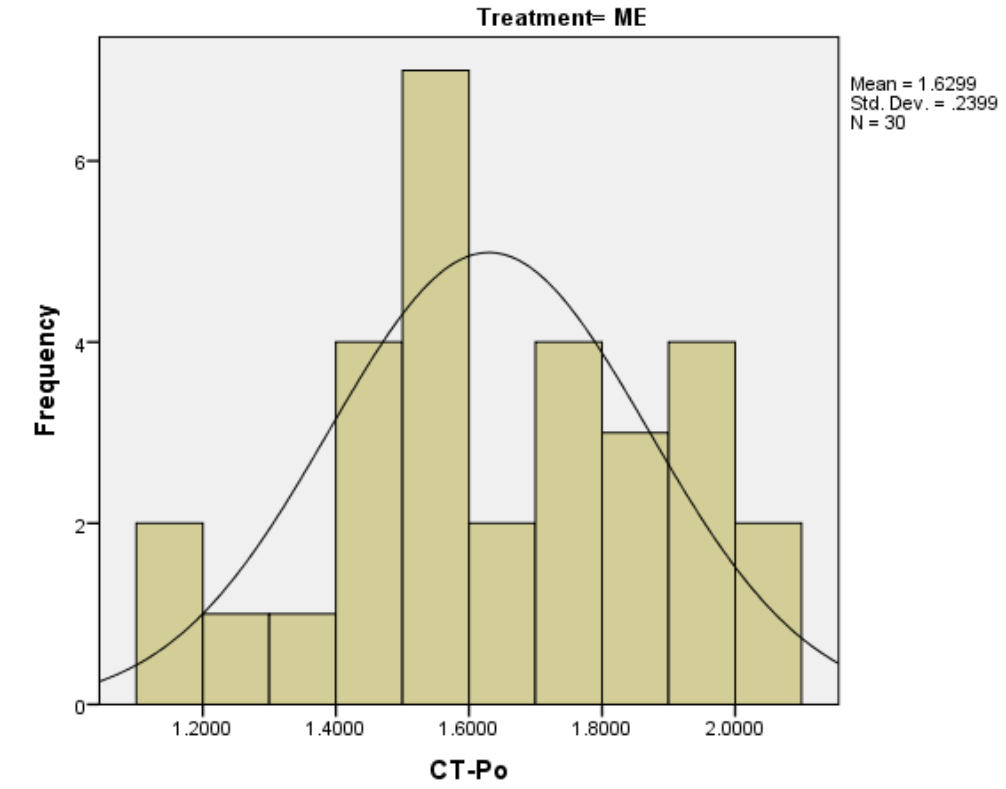


Figure 4.61. Posttest of C/T in the ME Group

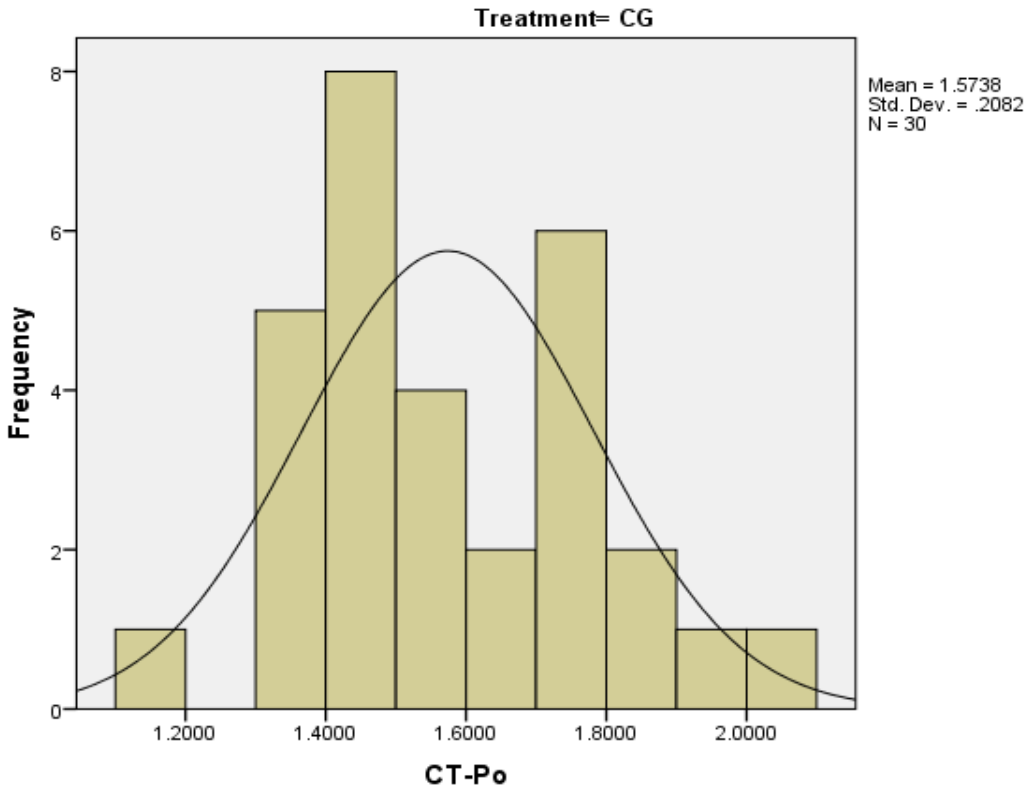


Figure 4.62. Posttest of C/T in the NF Group

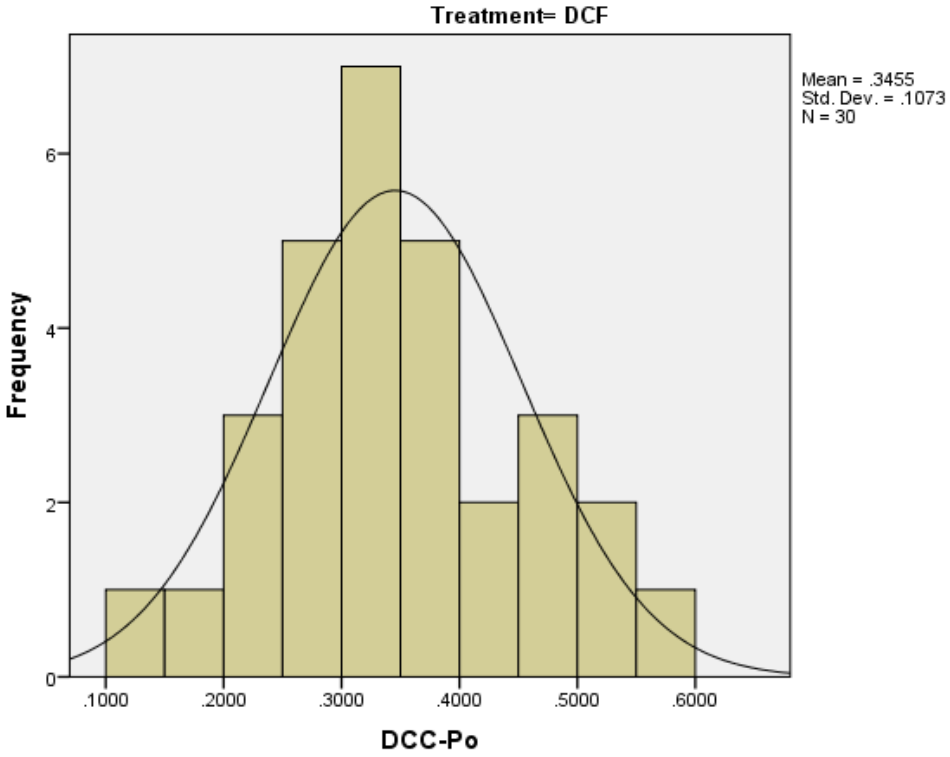


Figure 4.63. Posttest of DC/C in the DCF Group

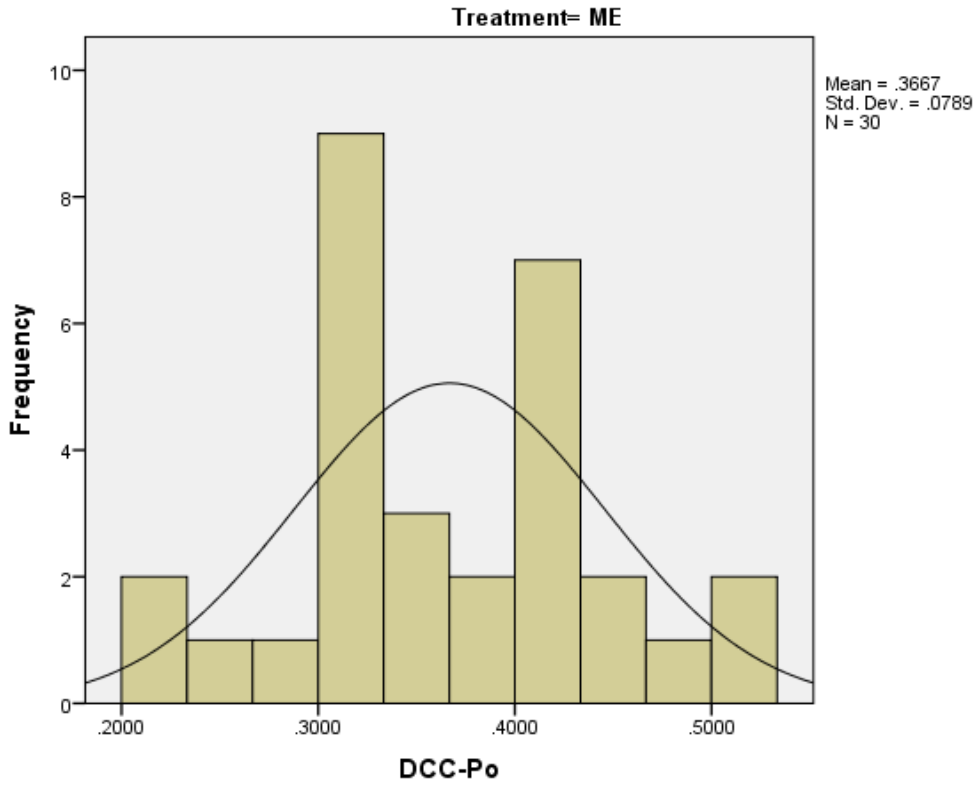


Figure 4.64. Posttest of DC/C in the ME Group

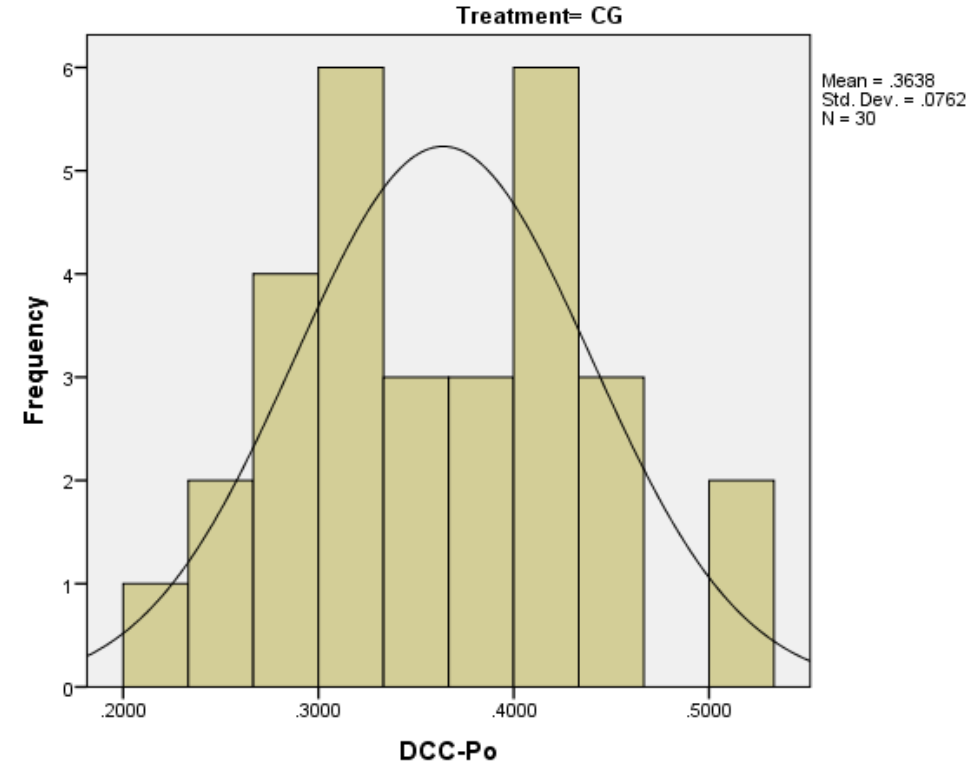


Figure 4.65. Posttest of DC/C in the NF Group

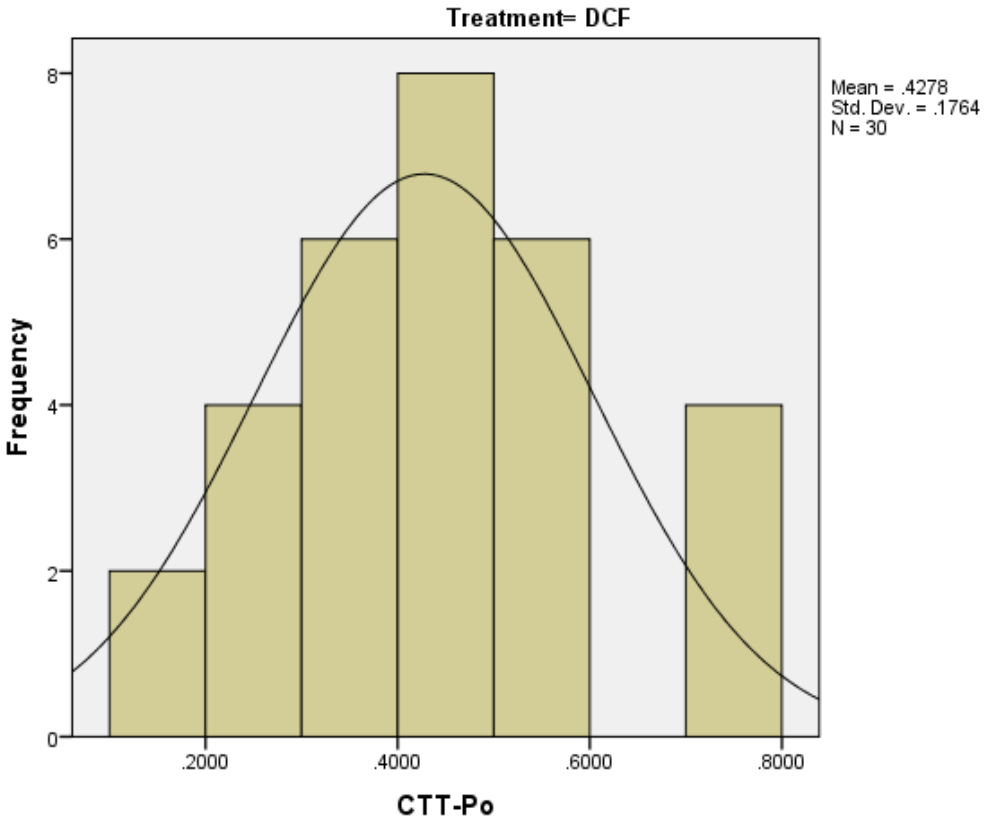


Figure 4.66. Posttest of CT/T in the DCF Group

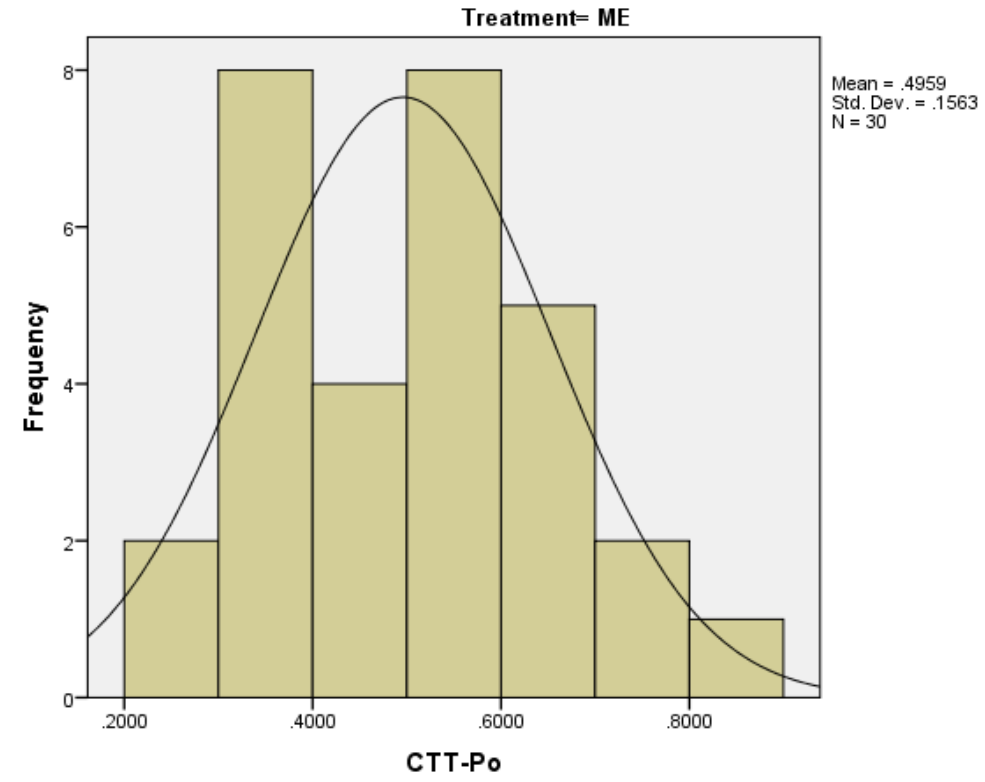


Figure 4.67. Posttest of CT/T in the ME Group

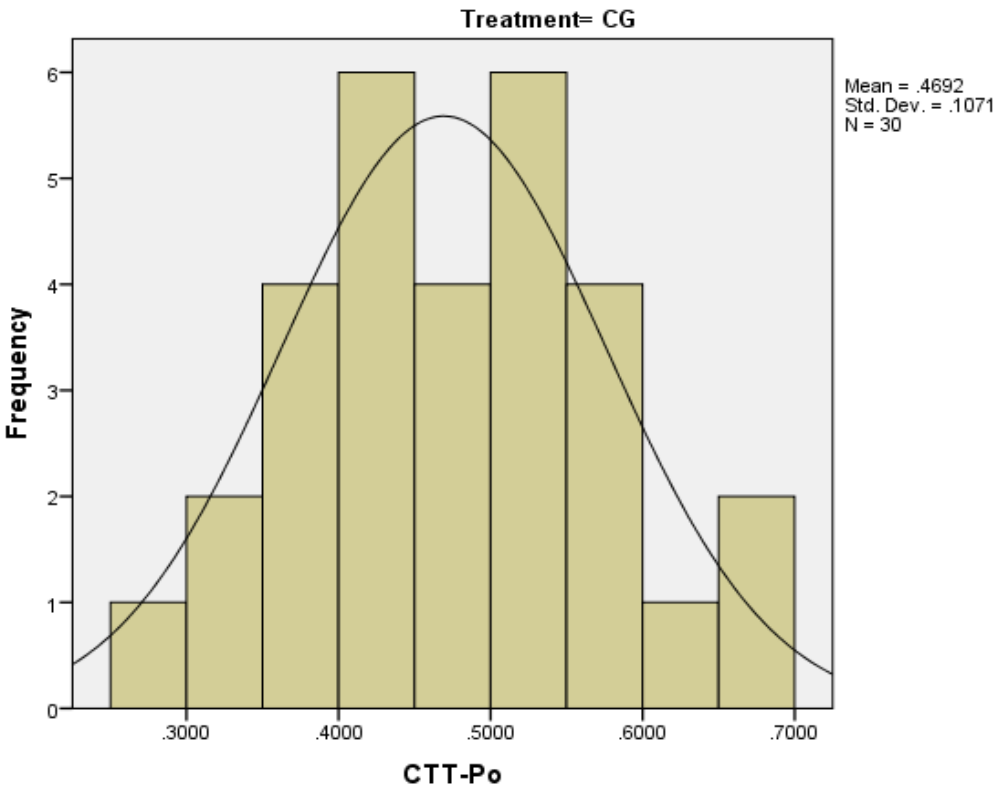


Figure 4.68. Posttest of CT/T in the NF Group

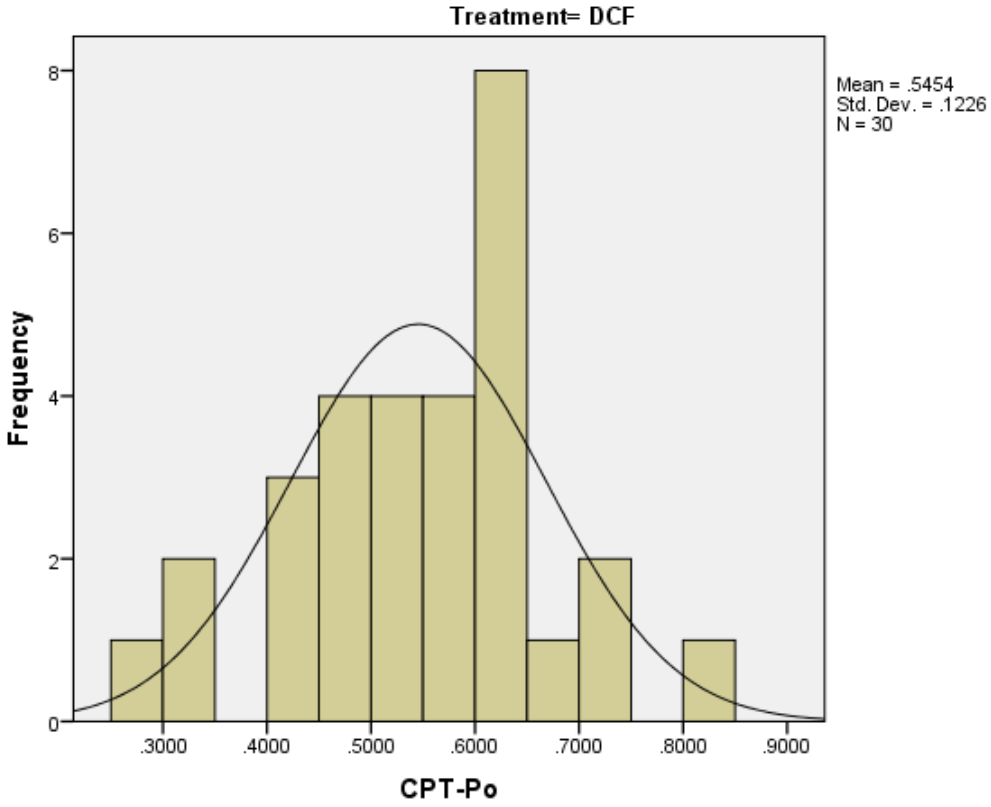


Figure 4.69. Posttest of CP/T in the DCF Group

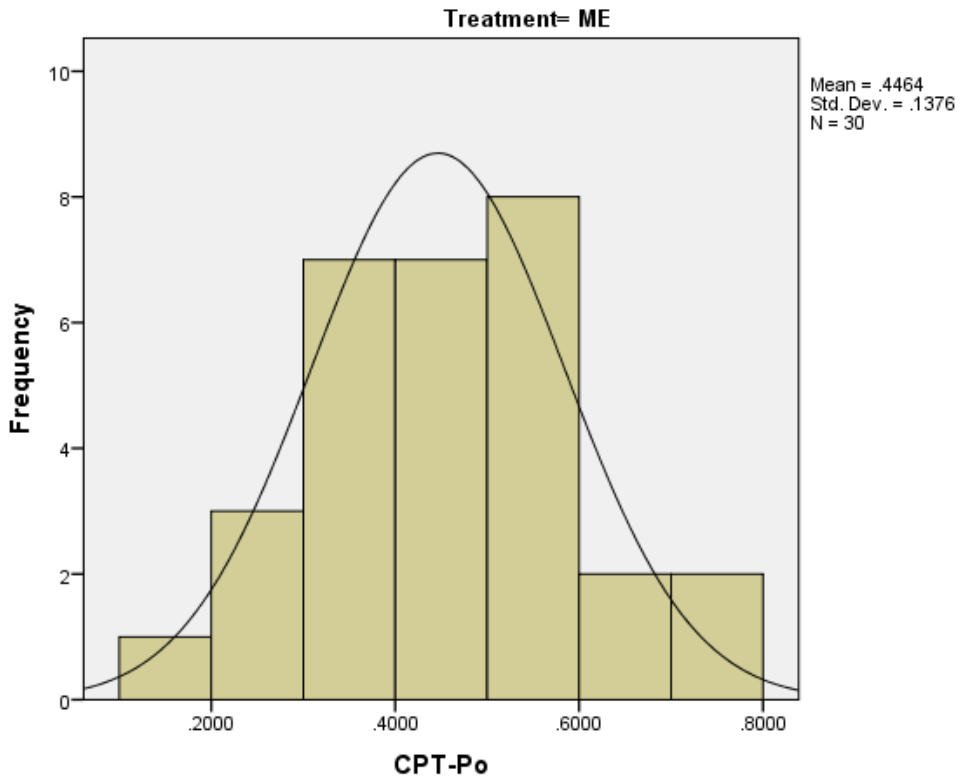


Figure 4.70. Posttest of CP/T in the ME Group

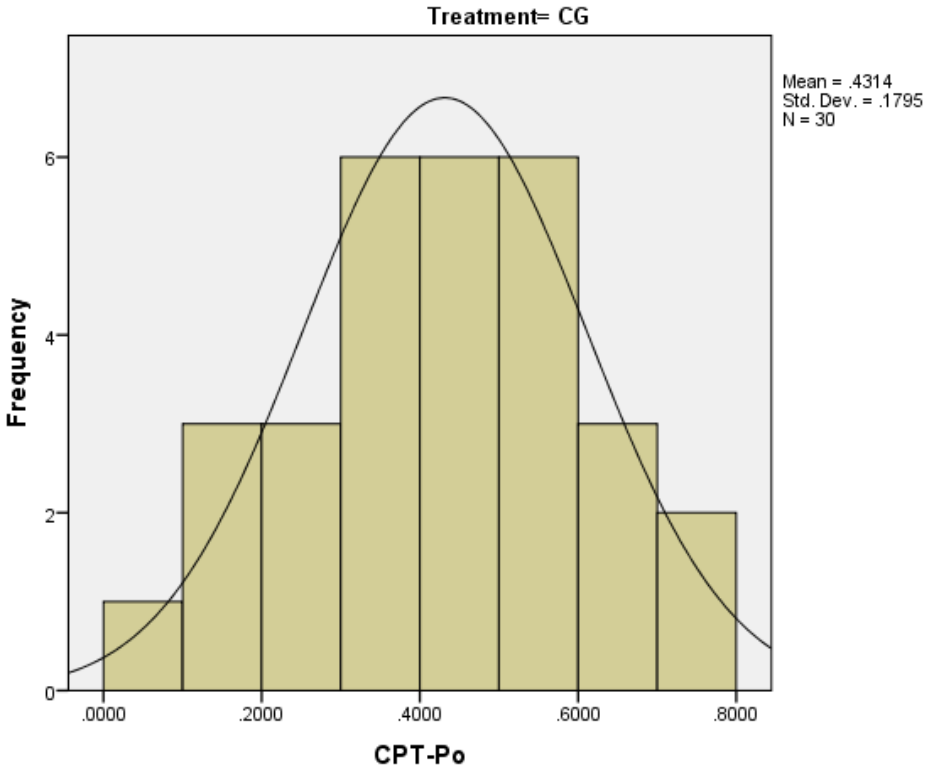


Figure 4.71. Posttest of CP/T in the NF Group

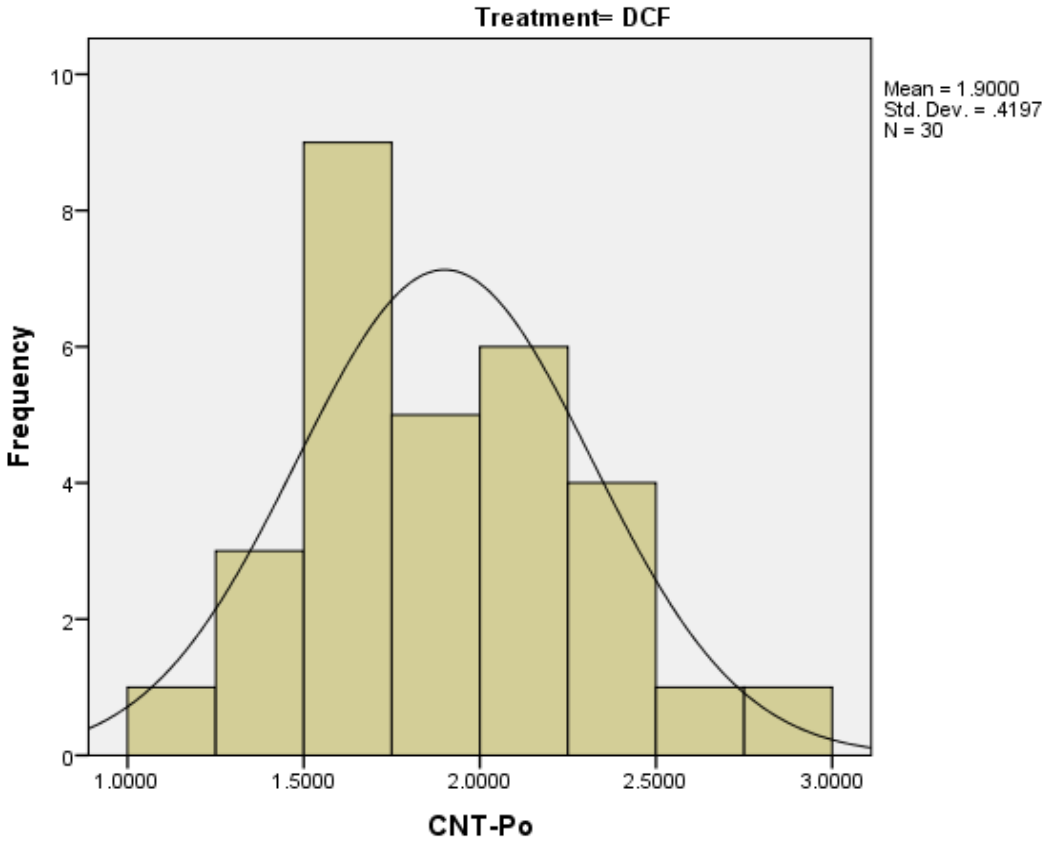


Figure 4.72. Posttest of CN/T in the DCF Group

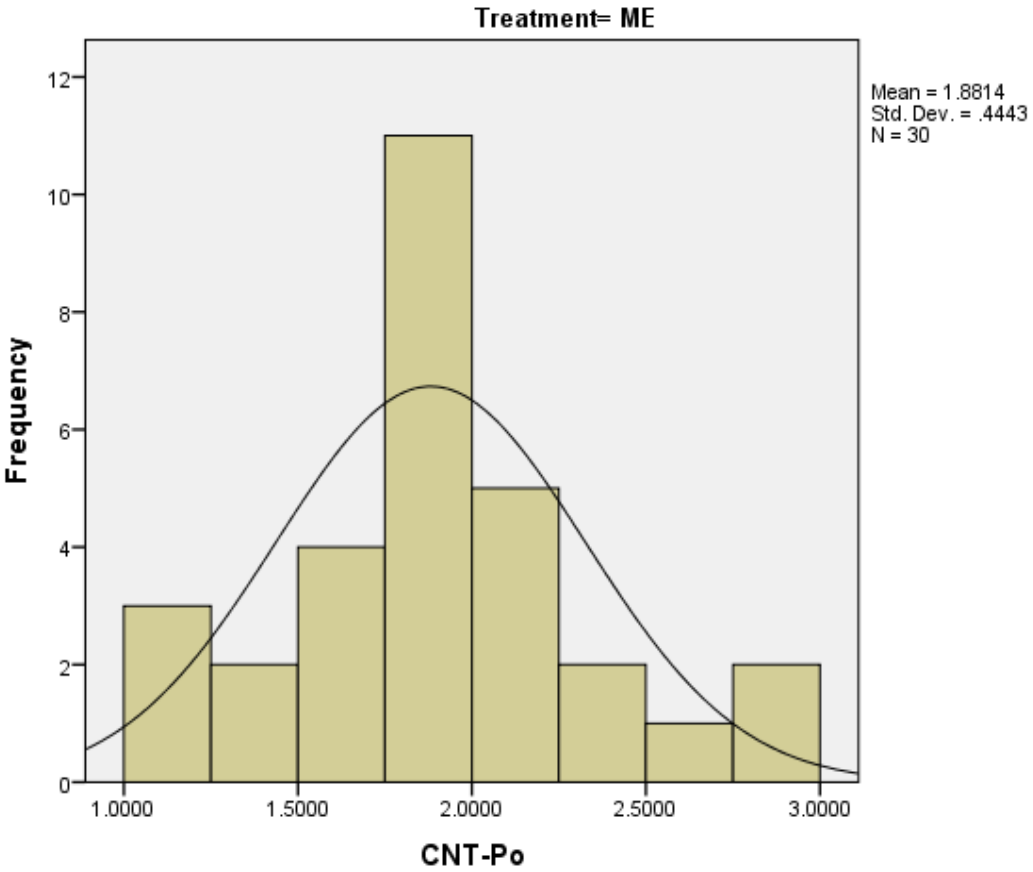


Figure 4.73. Posttest of CN/T in the ME Group

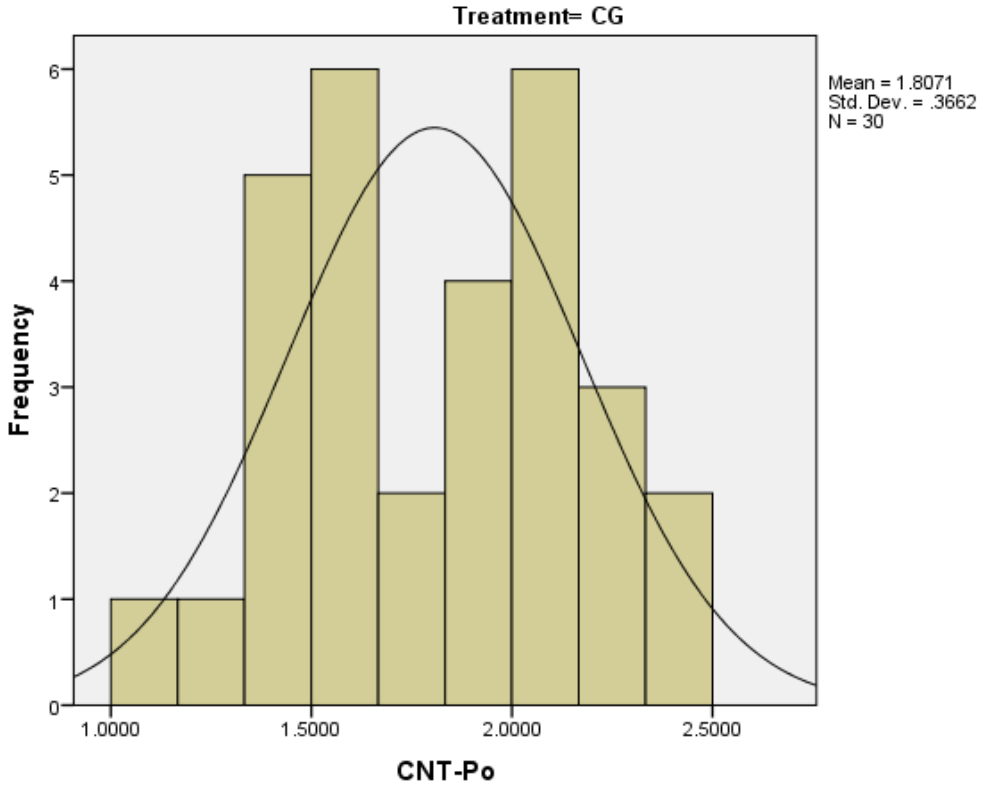


Figure 4.74. Posttest of CN/T in the NF Group

Appendix 8. Delayed-Posttest of Syntactic Complexity

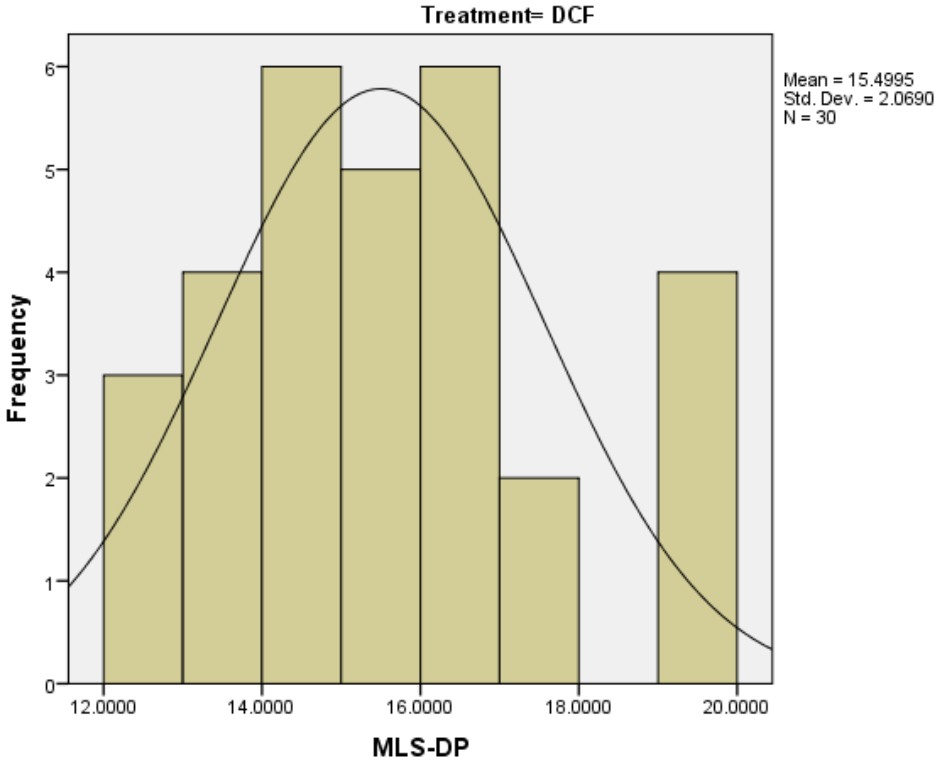


Figure 4.75. Delayed-Posttest of MLS in the DCF Group

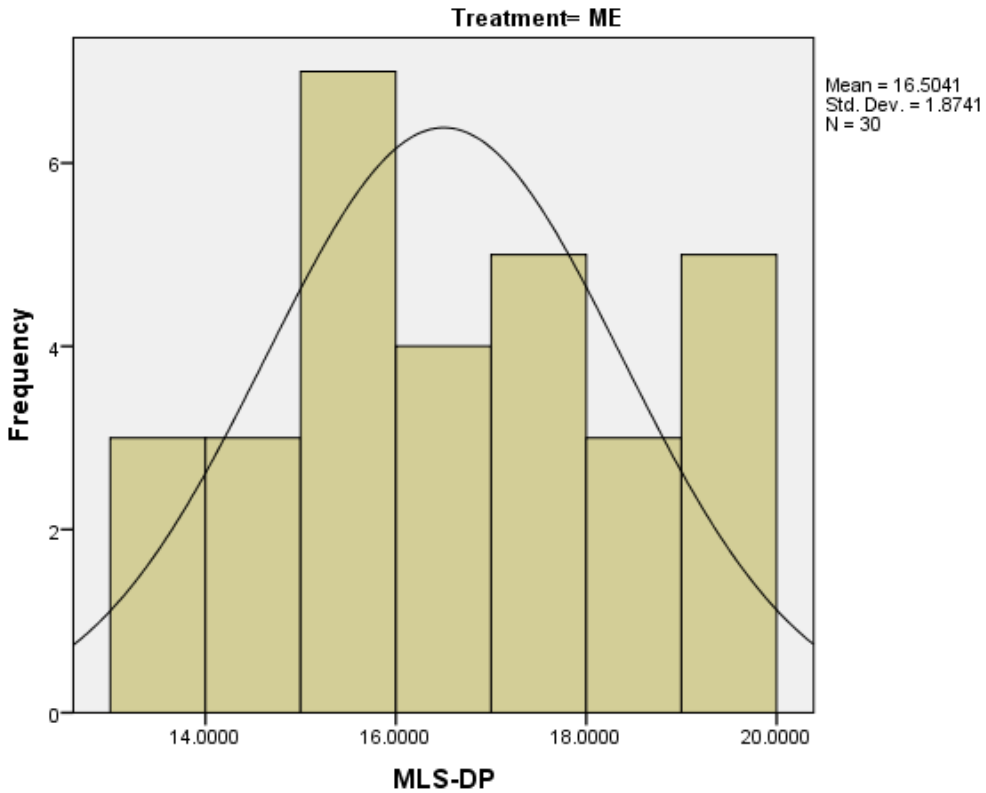


Figure 4.76. Delayed-Posttest of MLS in the ME Group

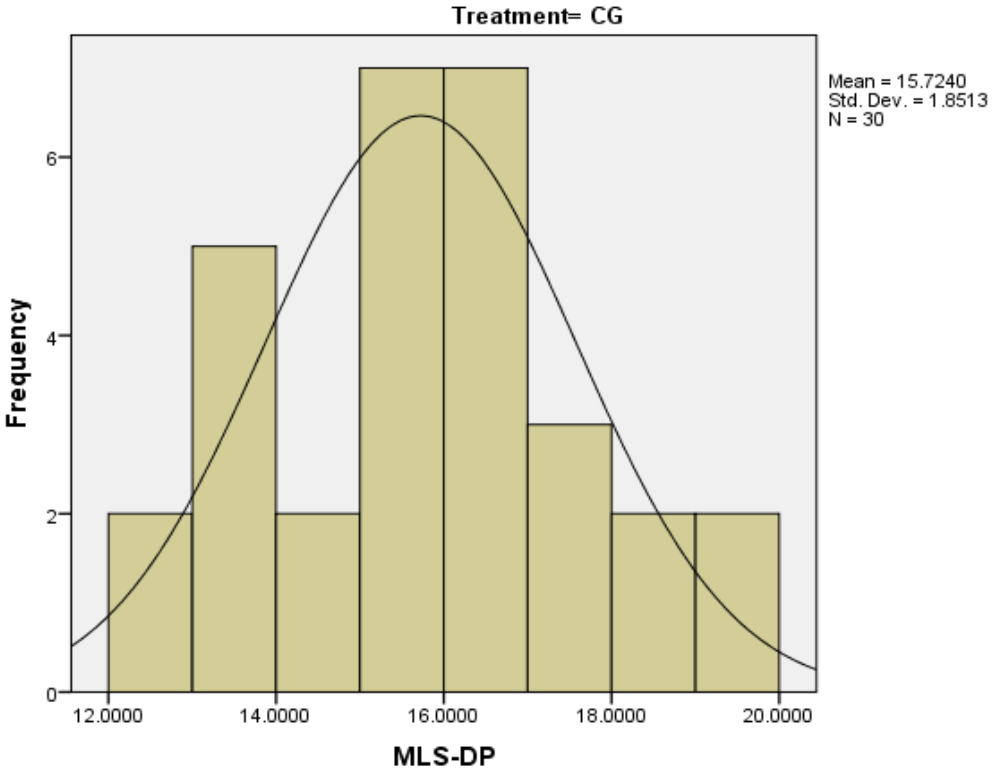


Figure 4.77. Delayed-Posttest of MLS in the NF Group

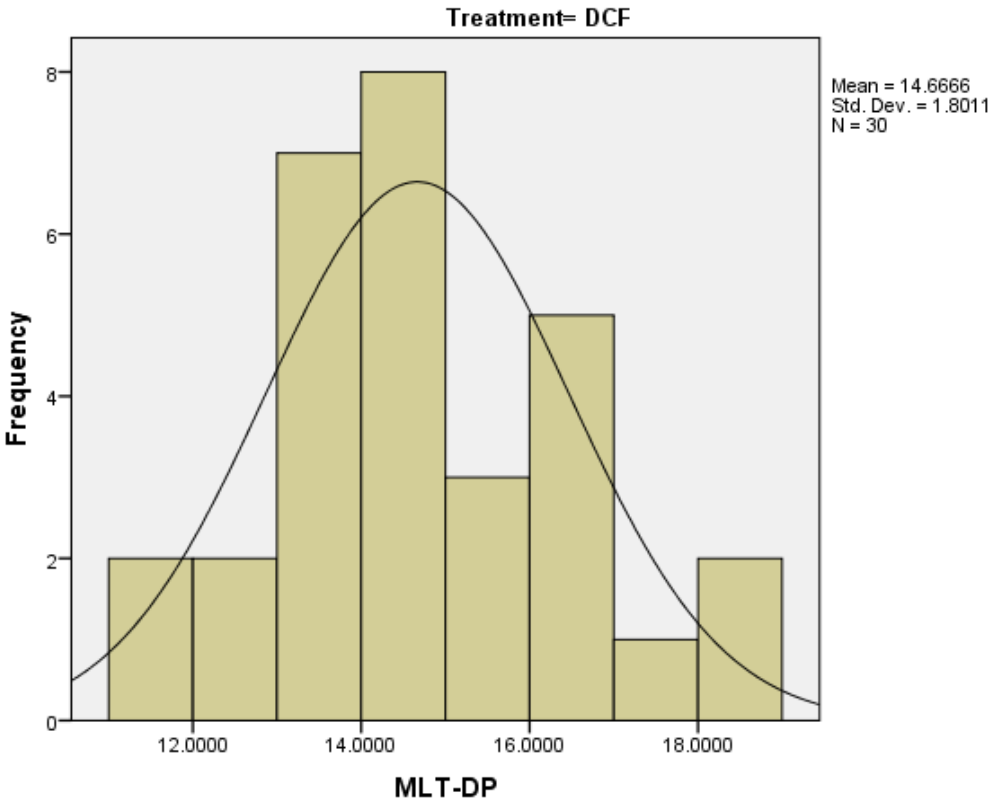


Figure 4.78. Delayed-Posttest of MLT in the DCF Group

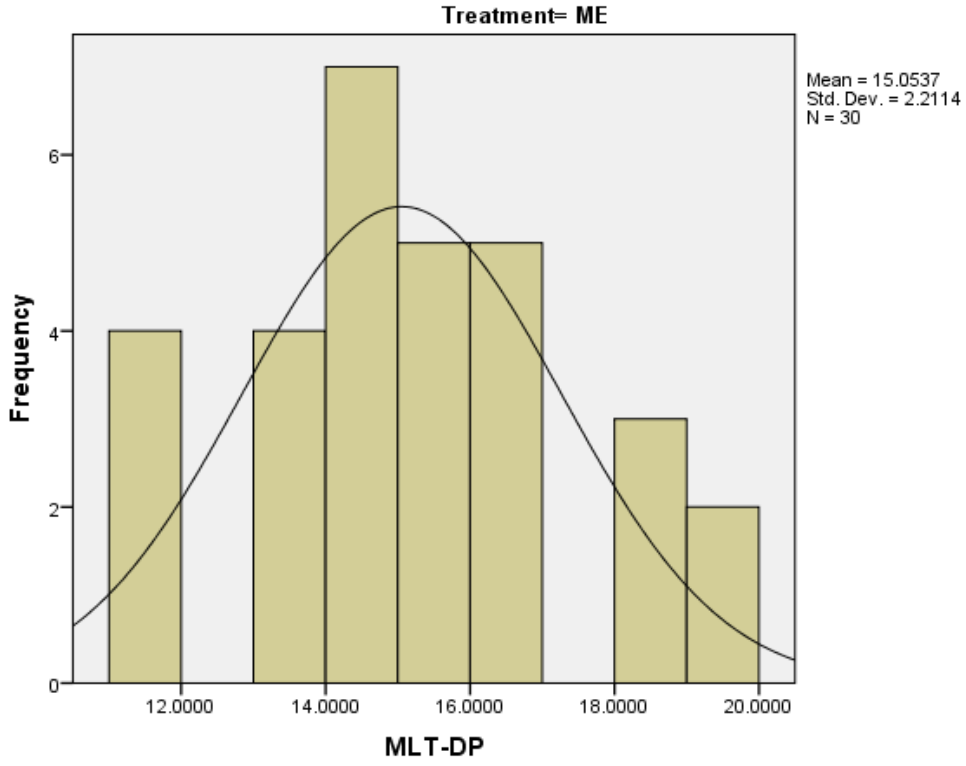


Figure 4.79. Delayed-Posttest of MLT in the ME Group

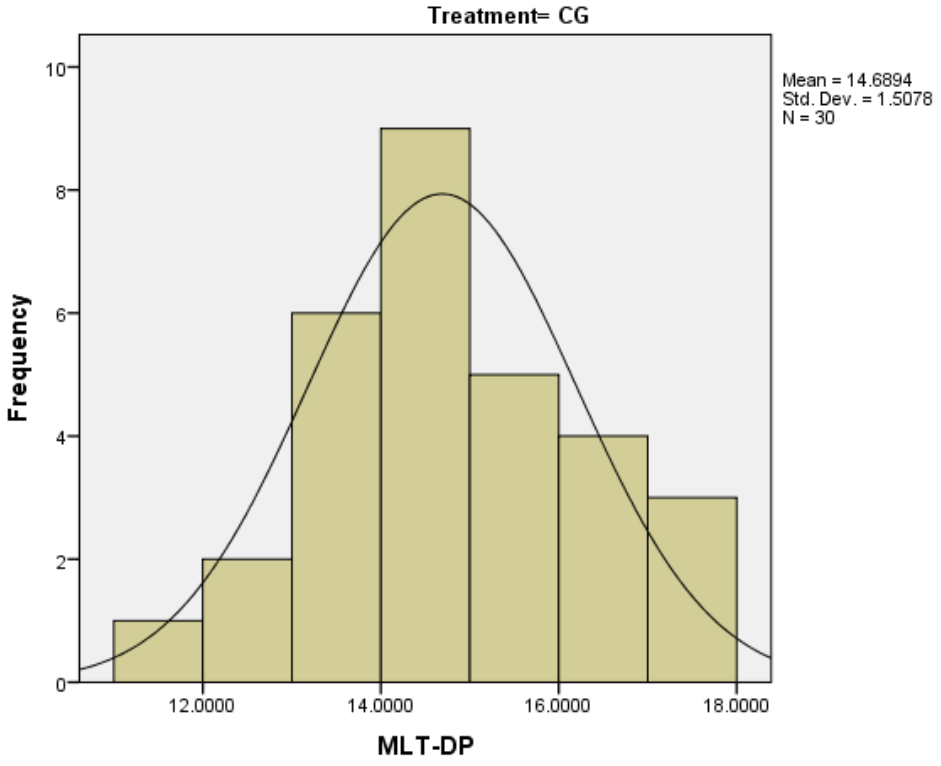


Figure 4.80. Delayed-Posttest of MLT in the NF Group

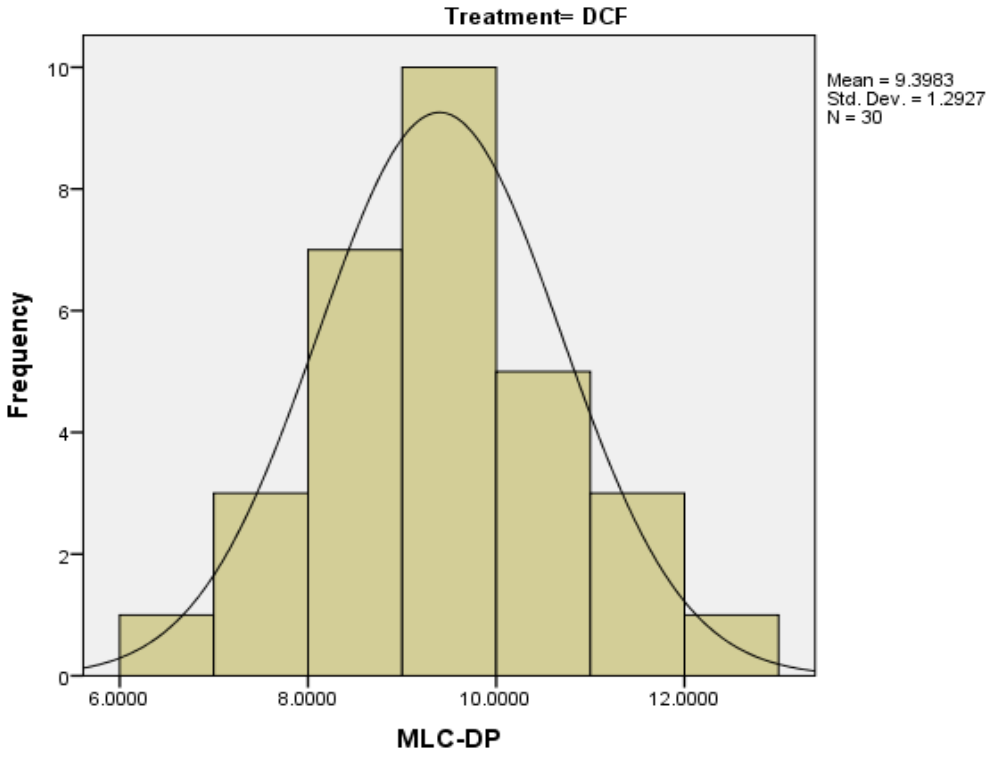


Figure 4.81. Delayed-Posttest of MLC in the DCF Group

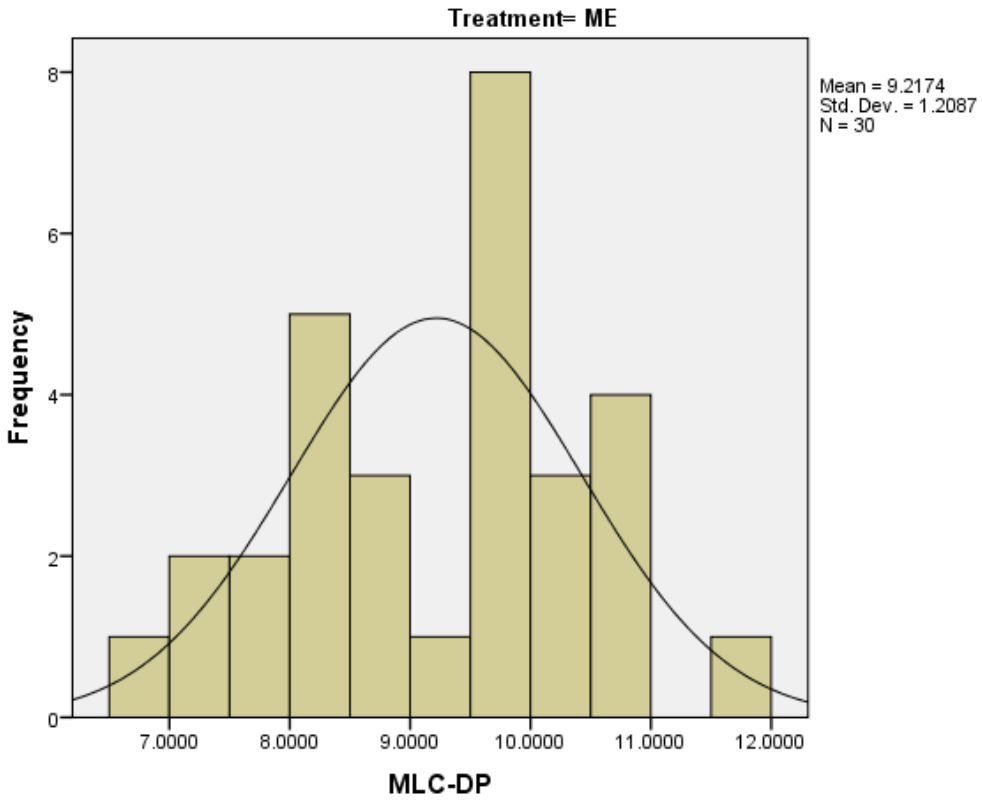


Figure 4.82. Delayed-Posttest of MLC in the ME Group

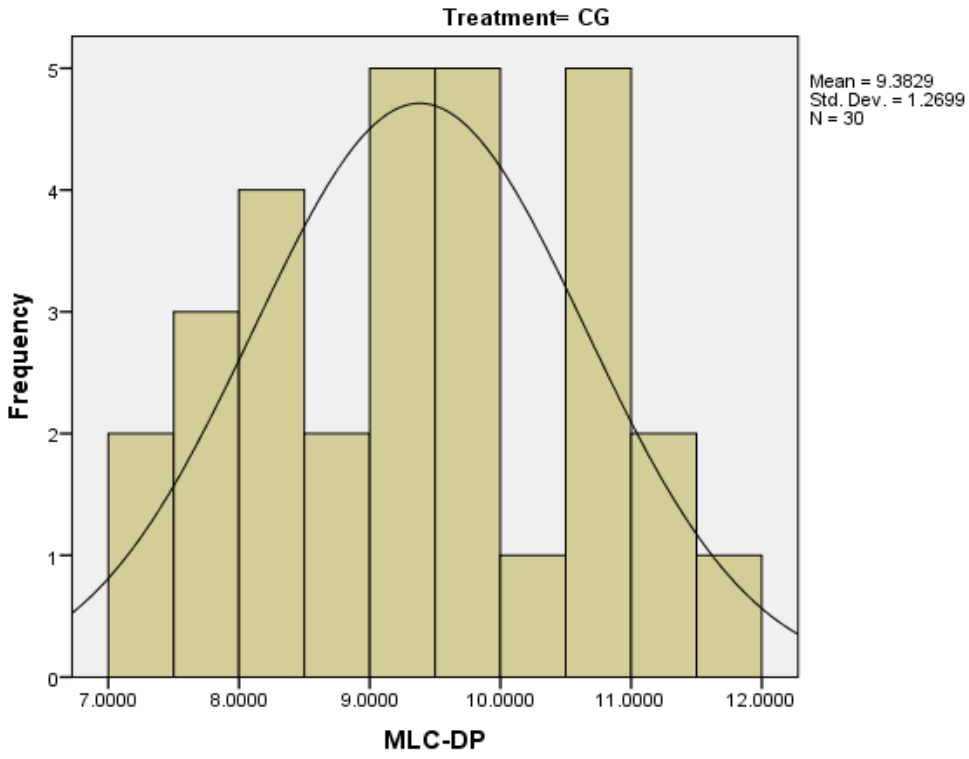


Figure 4.83. Delayed-Posttest of MLC in the NF Group

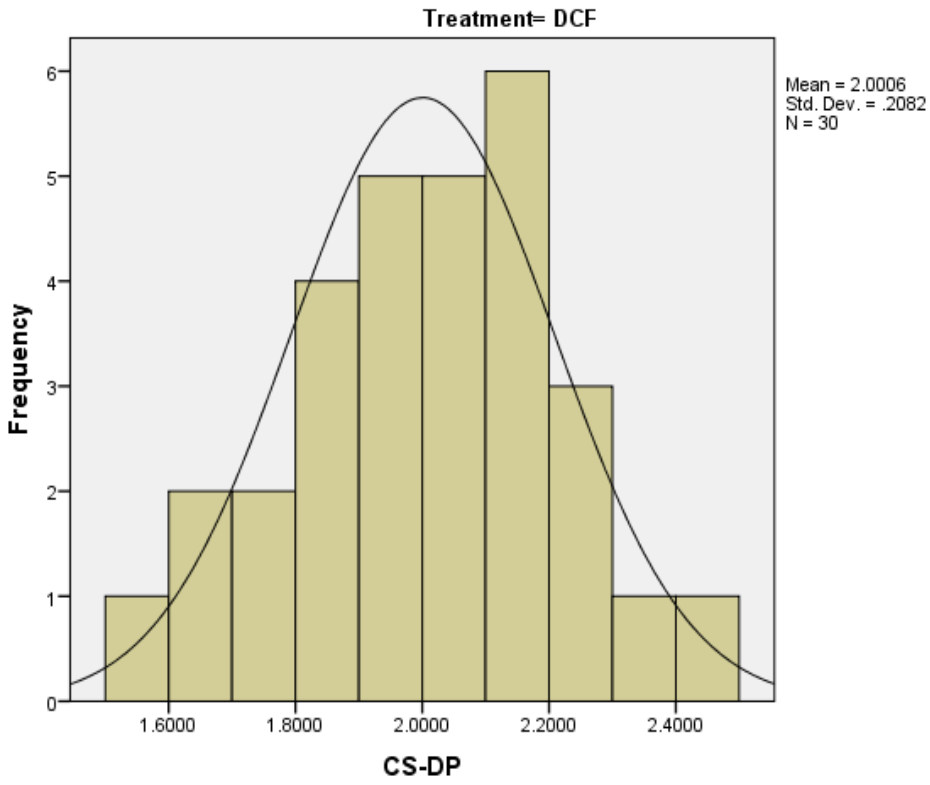


Figure 4.84. Delayed-Posttest of C/S in the DCF Group

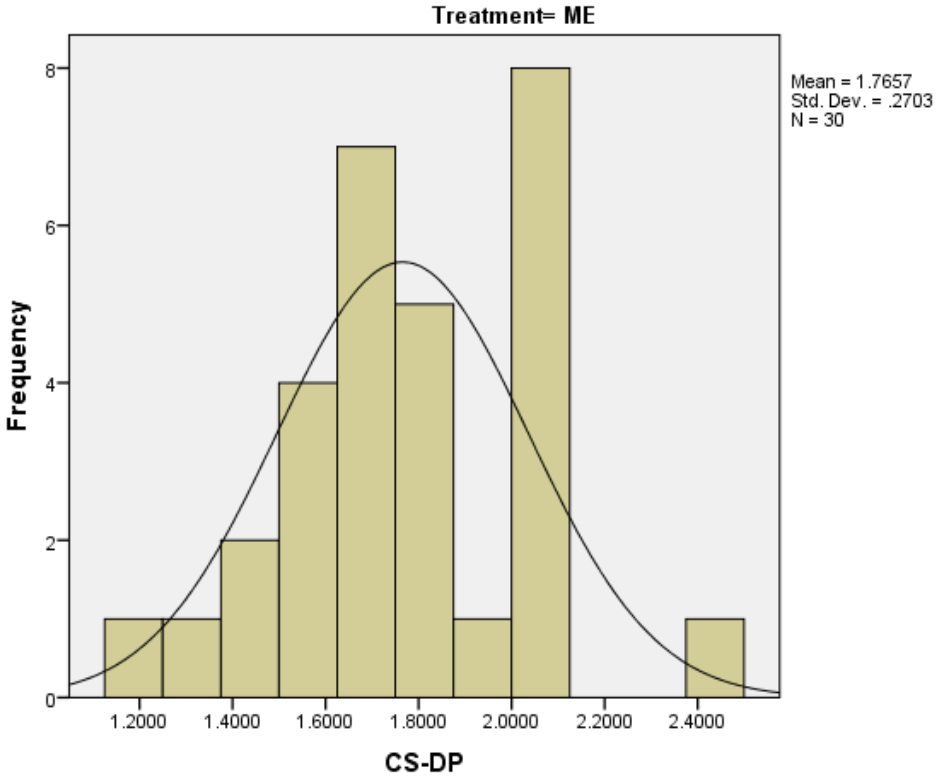


Figure 4.85. Delayed-Posttest of C/S in the ME Group

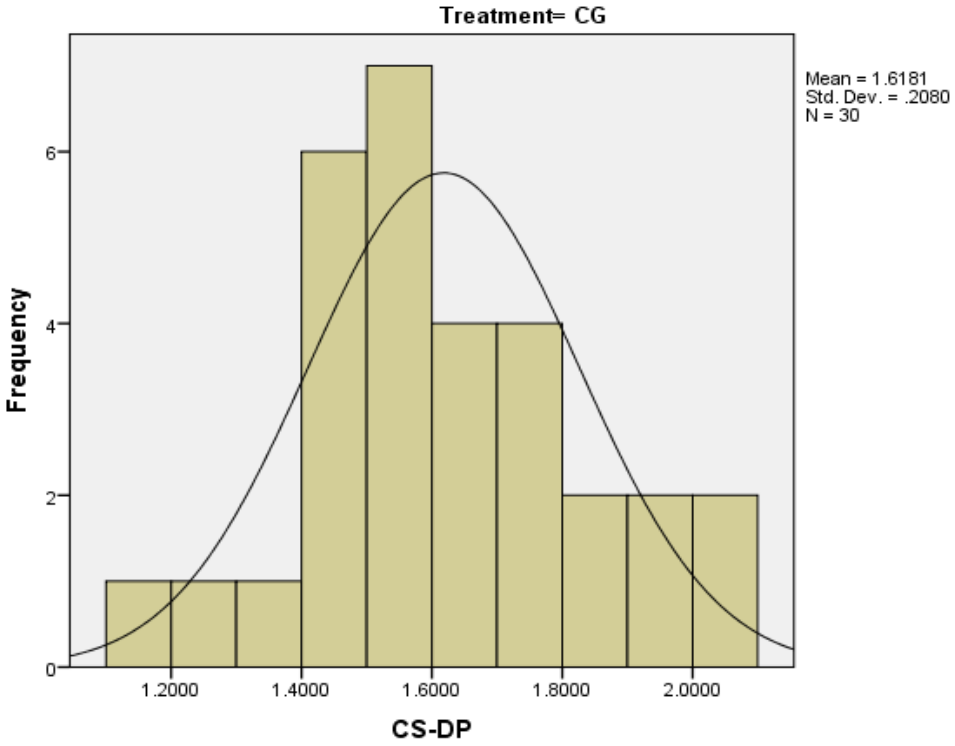


Figure 4.86. Delayed-Posttest of C/S in the NF Group

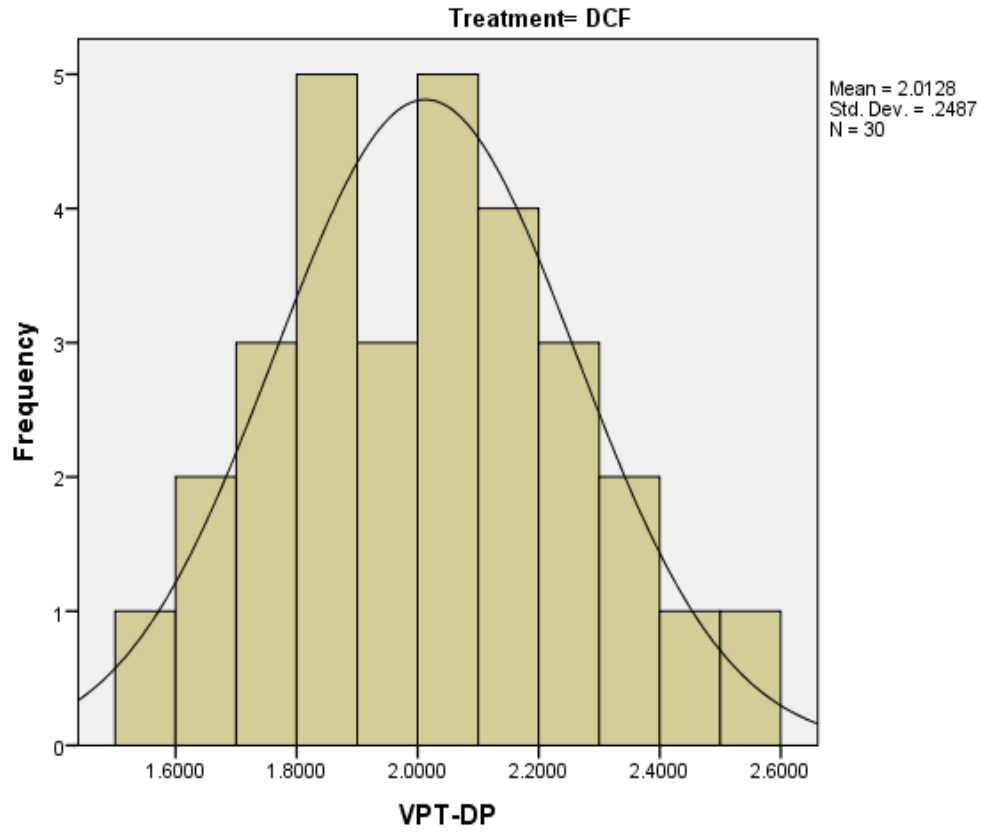


Figure 4.87. Delayed-Posttest of VP/T in the DCF Group

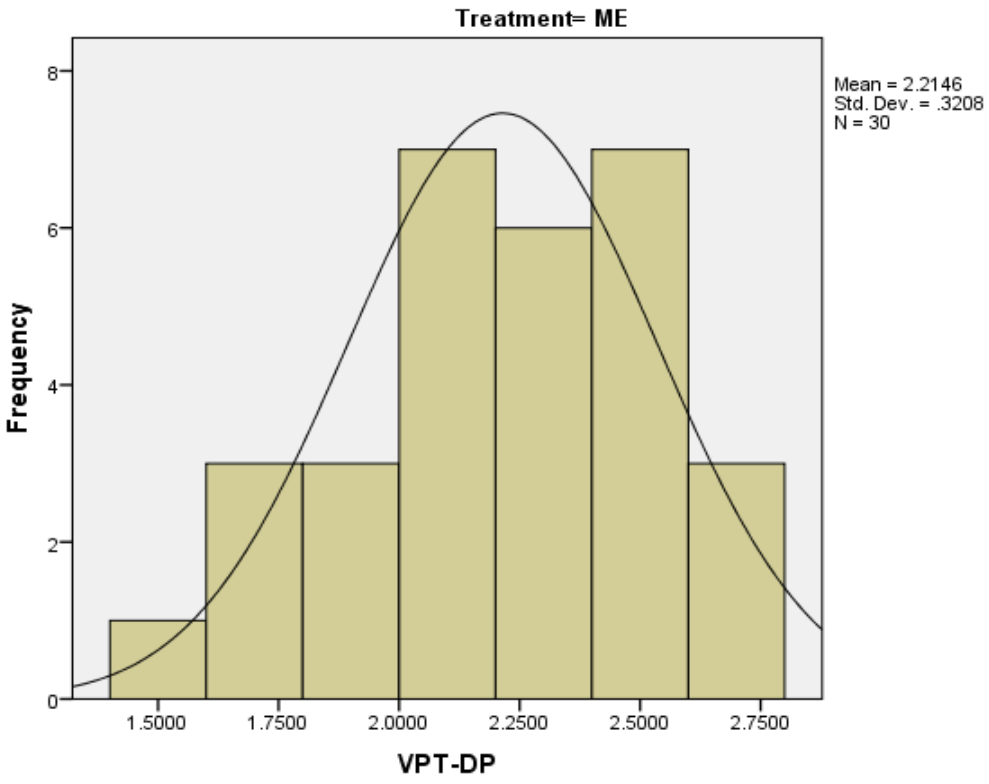


Figure 4.88. Delayed-Posttest of VP/T in the ME Group

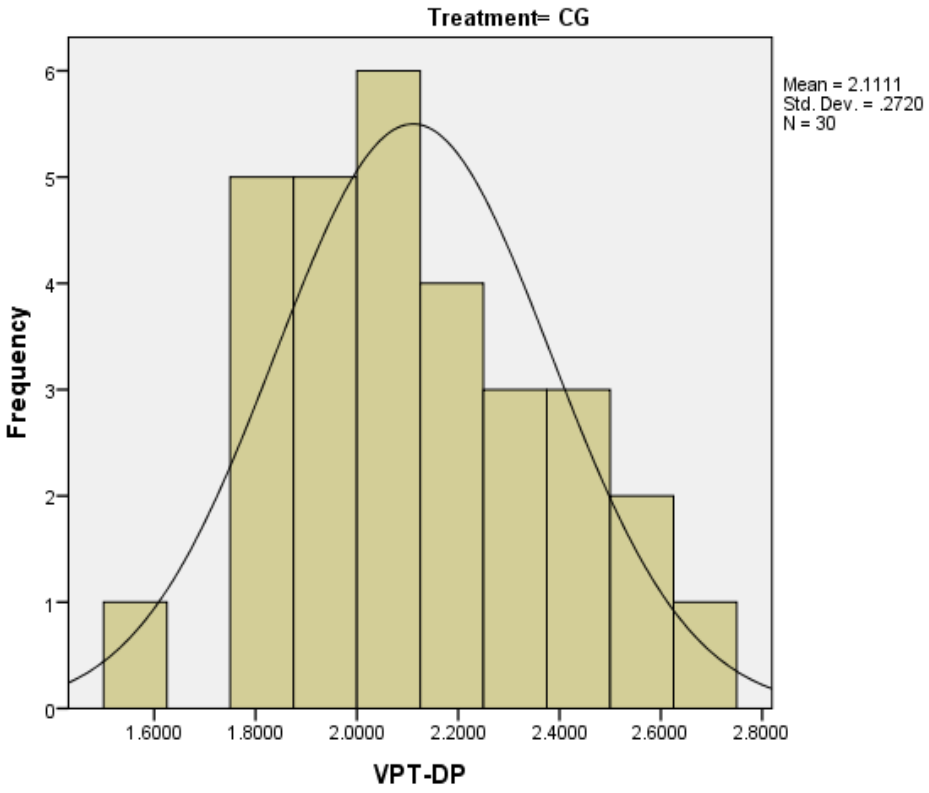


Figure 4.89. Delayed-Posttest of VP/T in the NF Group

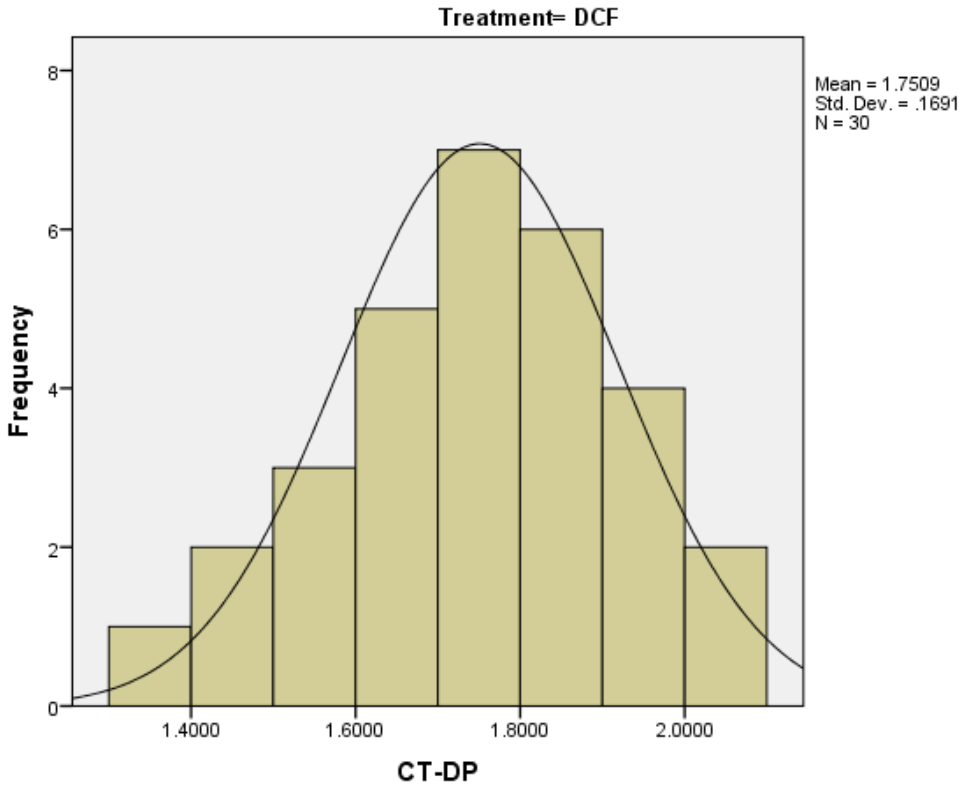


Figure 4.90. Delayed-Posttest of C/T in the DCF Group

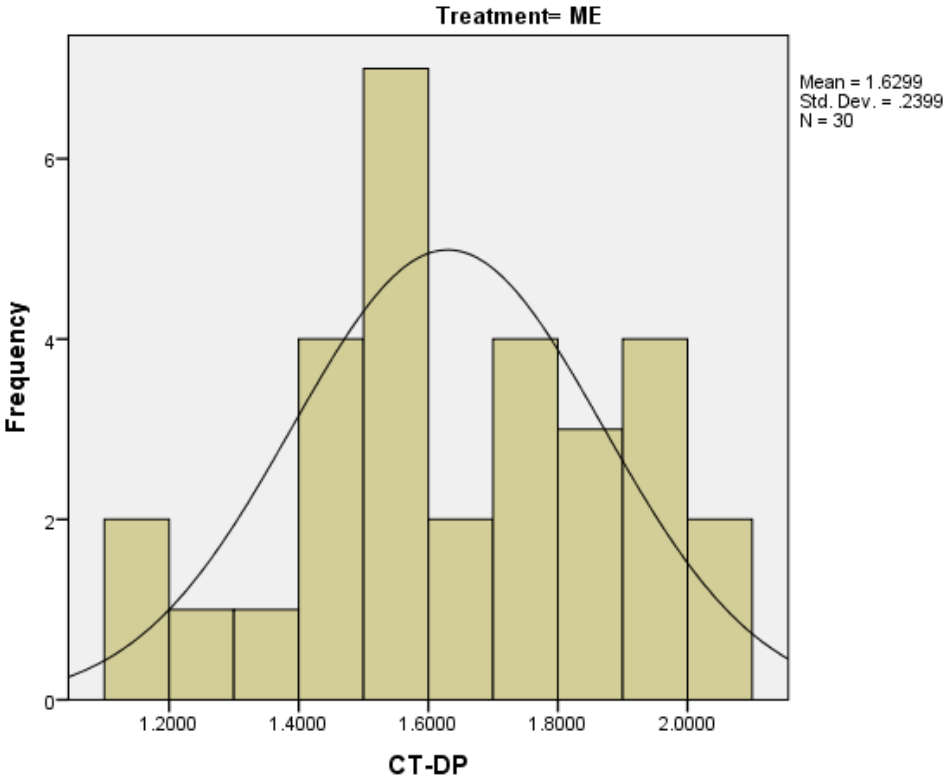


Figure 4.91. Delayed-Posttest of C/T in the ME Group

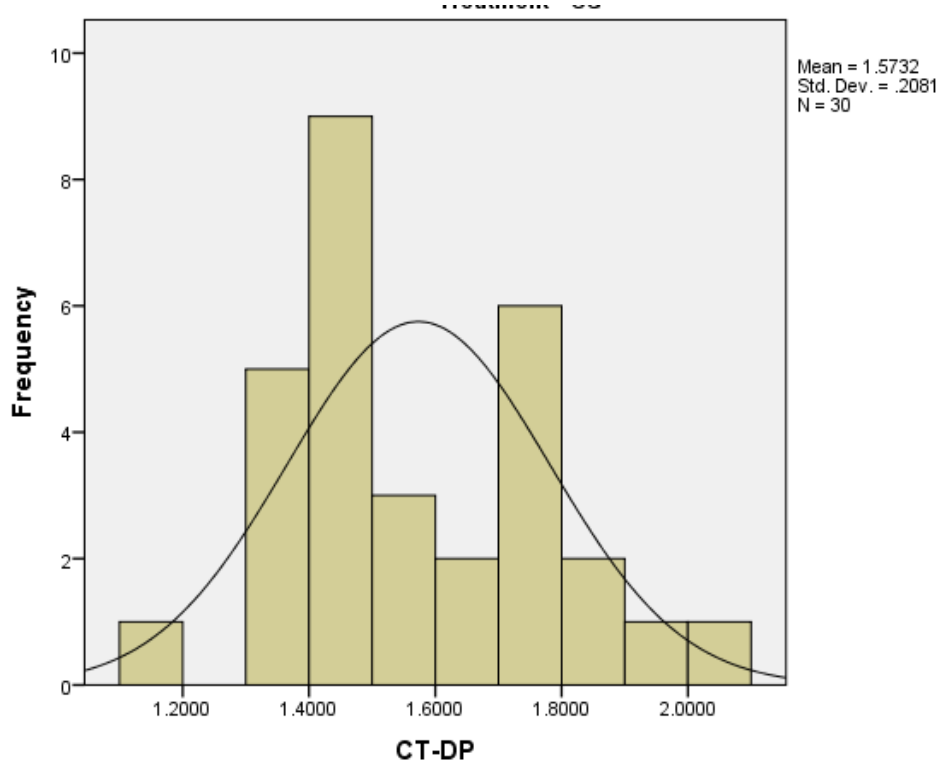


Figure 4.92. Delayed-Posttest of C/T in the NF Group

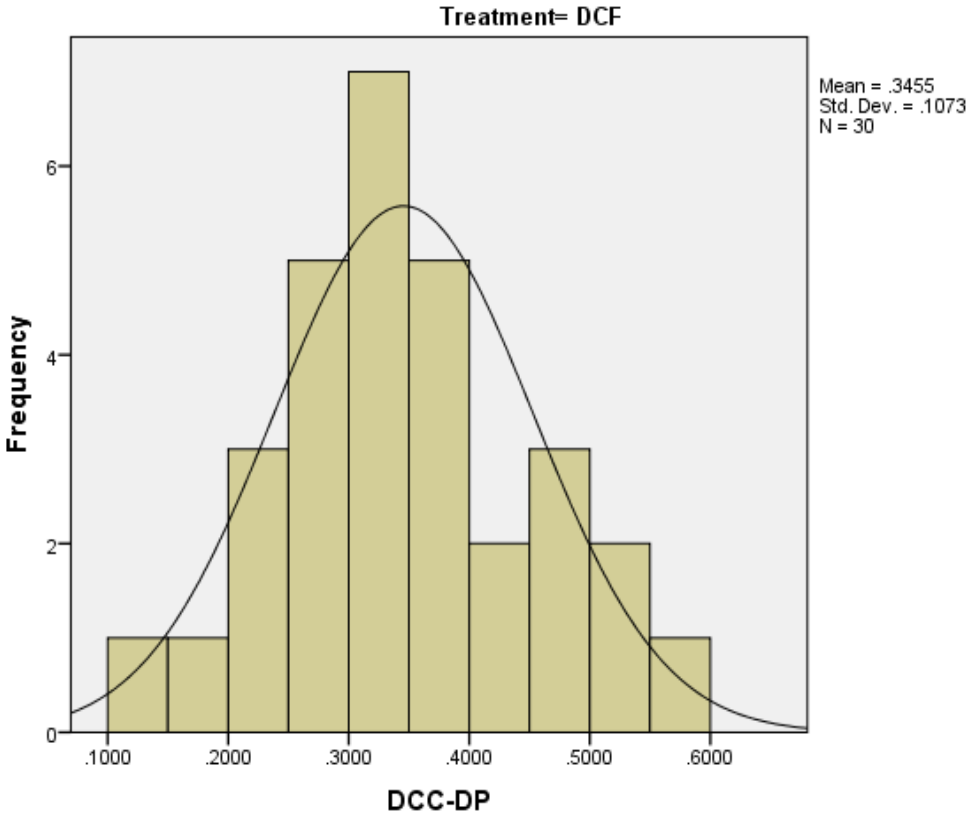


Figure 4.93. Delayed-Posttest of DC/C in the DCF Group

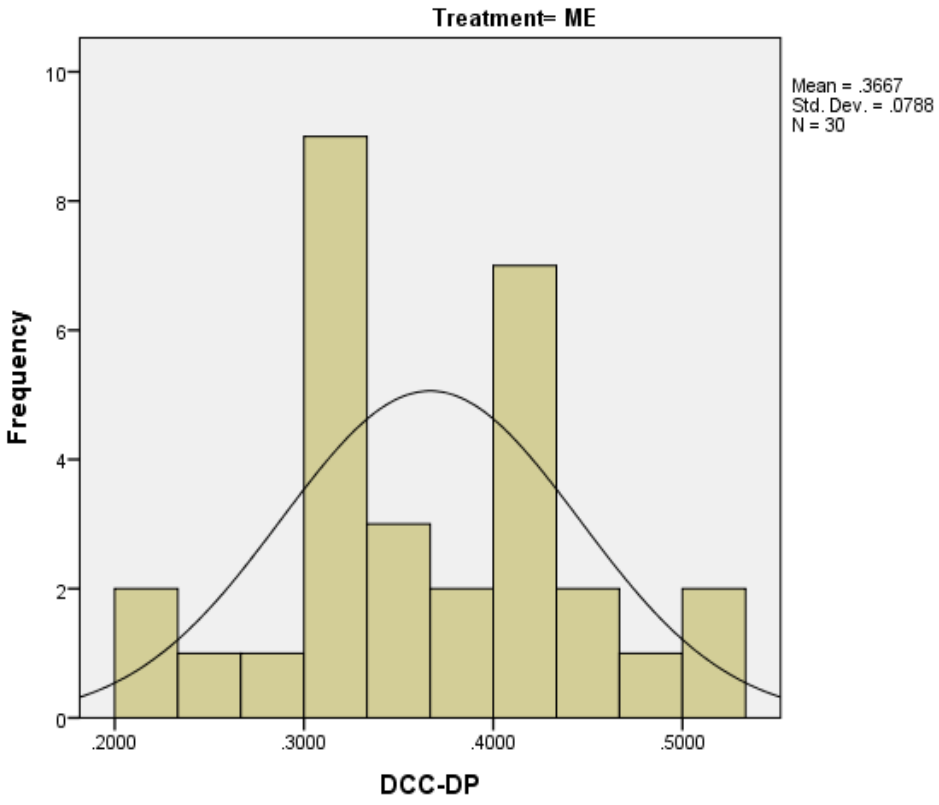


Figure 4.94. Delayed-Posttest of DC/C in the ME Group

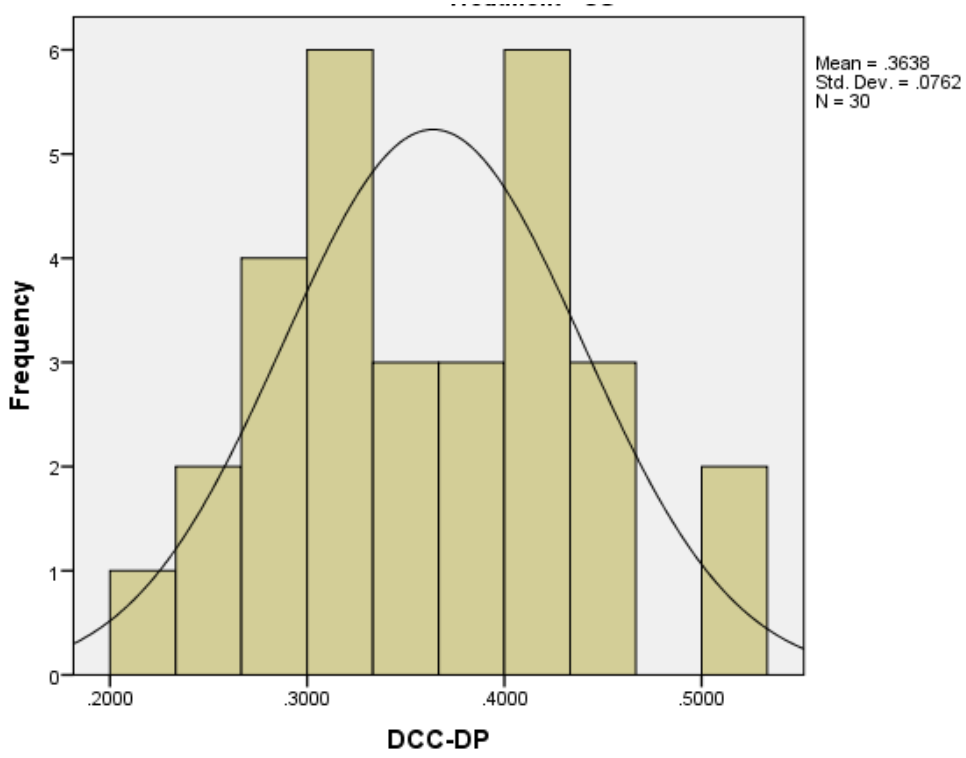


Figure 4.95. Delayed-Posttest of DC/C in the NF Group

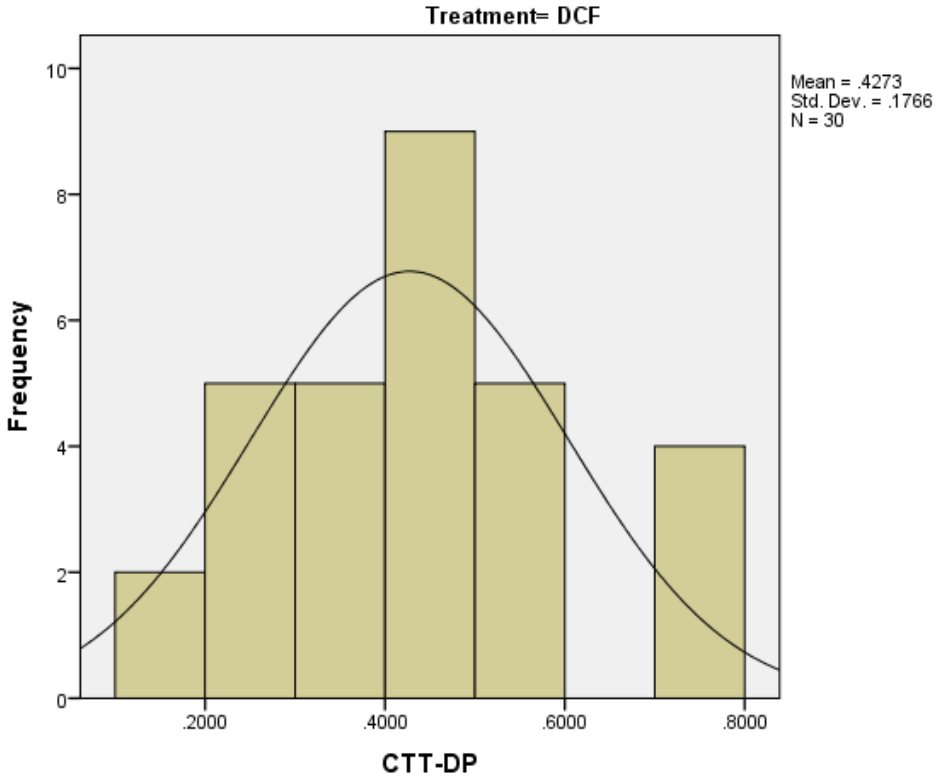


Figure 4.96. Delayed-Posttest of CT/T in the DCF Group

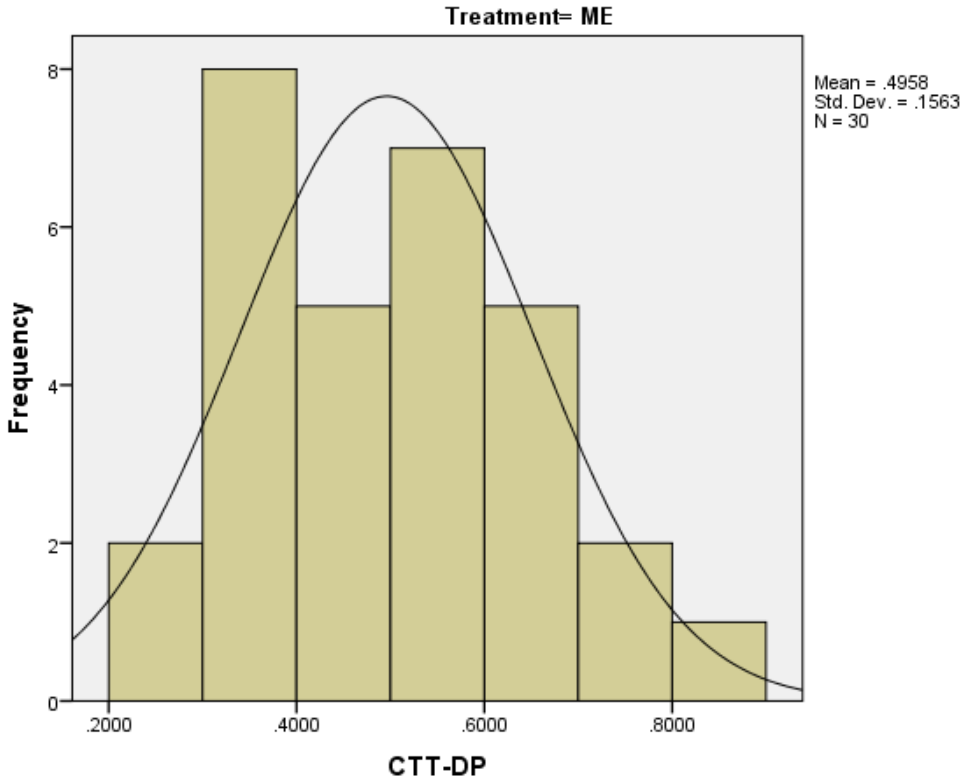


Figure 4.97. Delayed-Posttest of CT/T in the ME Group

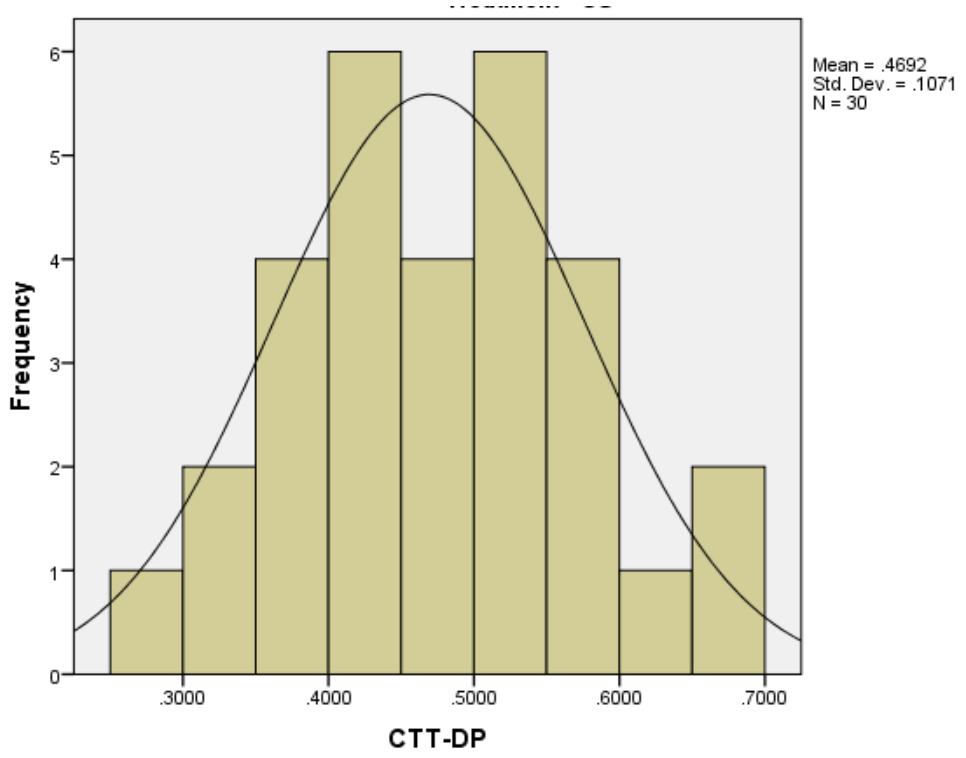


Figure 4.98. Delayed-Posttest of CT/T in the NF Group

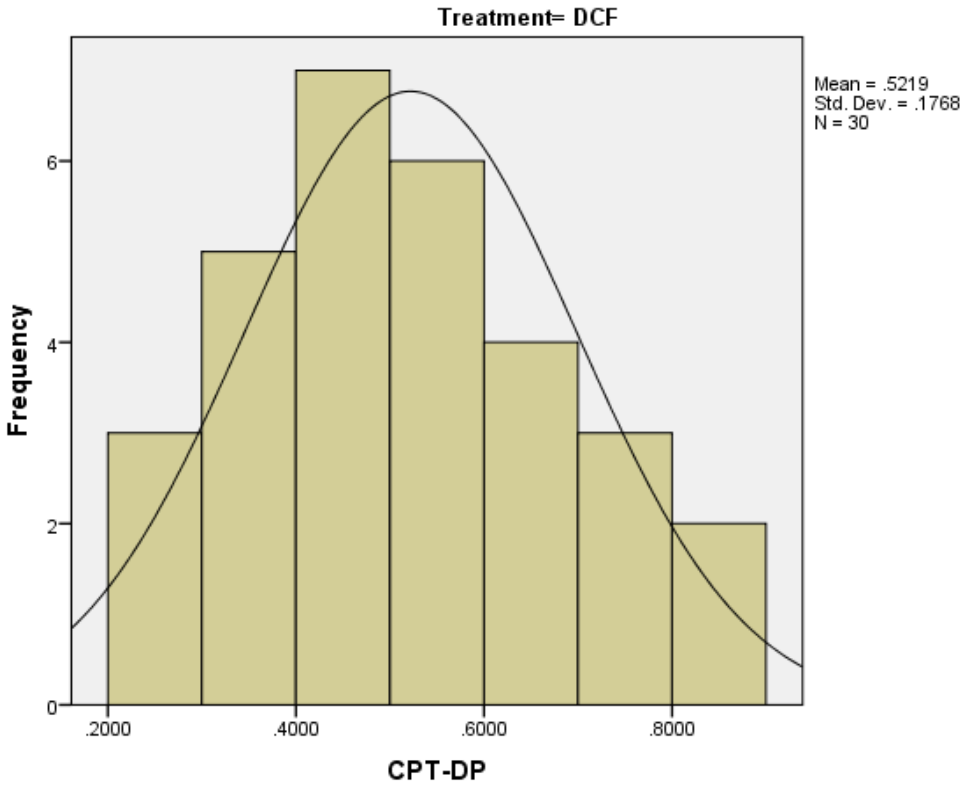


Figure 4.99. Delayed-Posttest of CP/T in the DCF Group

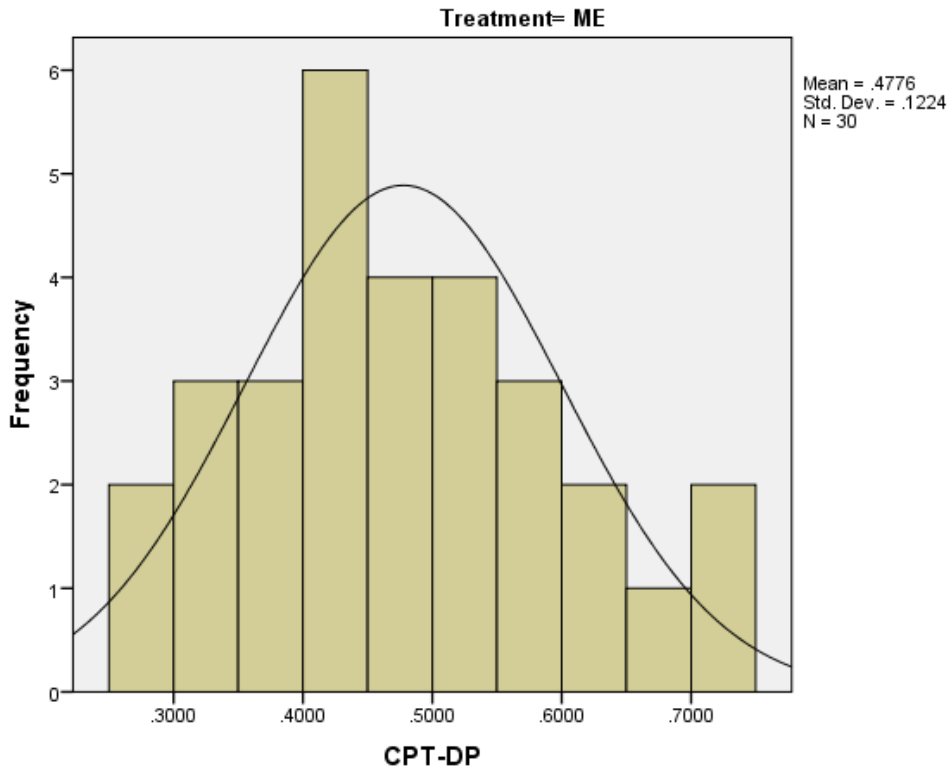


Figure 4.100. Delayed-Posttest of CP/T in the ME Group

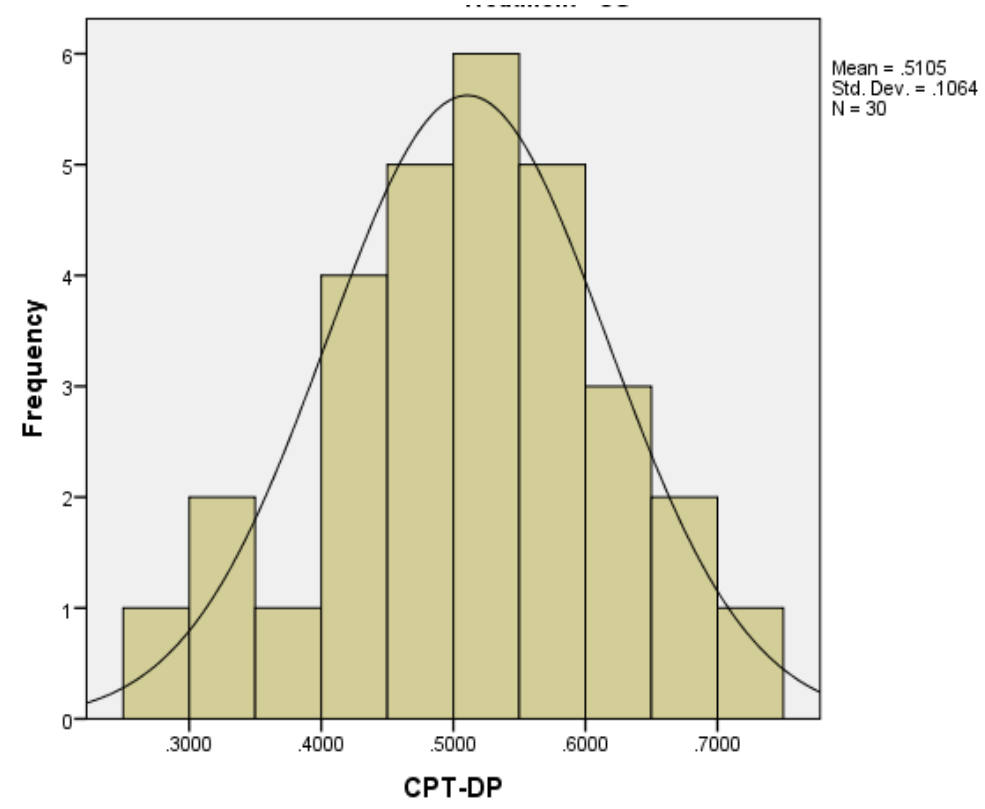


Figure 4.101. Delayed-Posttest of CP/T in the NF Group

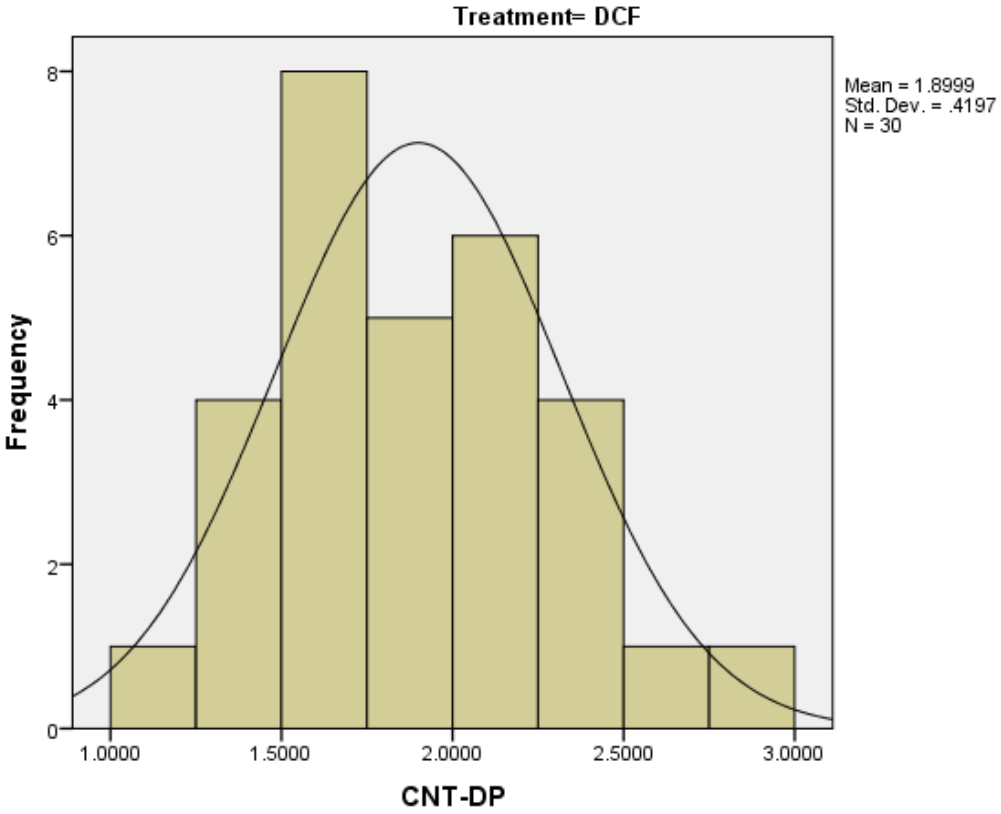


Figure 4.102. Delayed-Posttest of CN/T in the DCF Group

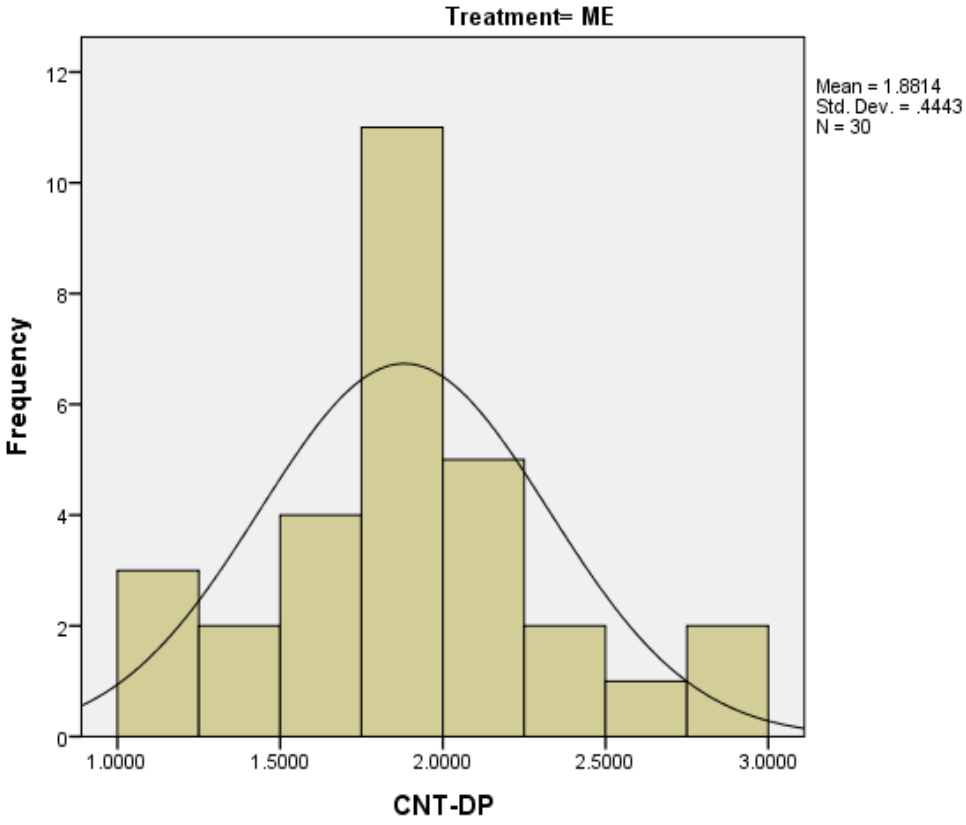


Figure 4.103. Delayed-Posttest of CN/T in the ME Group

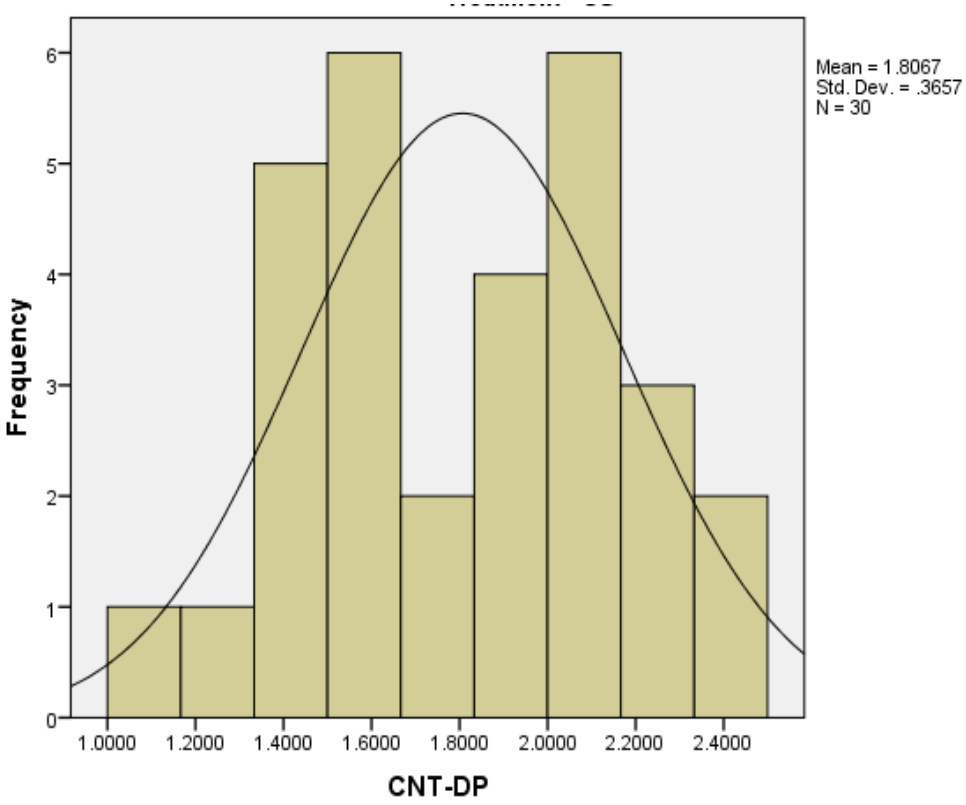


Figure 4.104. Delayed-Posttest of CN/T in the NF Group

Appendix 9. Normality Tests for QPT and Syntactic Accuracy

Table 4.50

Normality Tests for QPT and Syntactic Accuracy

Group		Skewness			Kurtosis		
		Statistic	Std. Error	Normality	Statistic	Std. Error	Normality
DCF	QPT	.05	.42	-.11	-.68	.83	-.81
	Pretest of Syntactic Accuracy	-.25	.42	-.59	-.26	.83	-.31
	Posttest of Syntactic Accuracy	-.41	.42	-.97	-.87	.83	-1.04
	Delayed-Posttest of Syntactic Accuracy	-.39	.42	-.92	-.14	.83	-.16
ME	QPT	-.06	.42	-.14	-.73	.83	-.87
	Pretest of Syntactic Accuracy	-.19	.42	-.45	-.87	.83	-1.04
	Posttest of Syntactic Accuracy	-.32	.42	-.76	-.38	.83	-.45
	Delayed-Posttest of Syntactic Accuracy	-.54	.42	-1.28	.09	.83	.10
NF	QPT	-.12	.42	-.28	-.60	.83	-.72
	Pretest of Syntactic Accuracy	-.12	.42	-.28	-.71	.83	-.85
	Posttest of Syntactic Accuracy	-.31	.42	-.73	-.83	.83	-1
	Delayed-Posttest of Syntactic Accuracy	-.25	.42	-.59	-.81	.83	-.97

Appendix 10. Normality Tests for Pretest of Syntactic Complexity

Table 4.51

Normality Tests for Pretest of Syntactic Complexity

Group		Skewness			Kurtosis		
		Statistic	Std. Error	Normality	Statistic	Std. Error	Normality
DCF	MLS	.29	.42	.69	-.96	.83	-1.15
	MLT	-.08	.42	-.19	-.17	.83	-.20
	MLC	-.33	.42	-.78	-.25	.83	-.30
	C/S	.08	.42	.19	-.61	.83	-.73
	VP/T	.13	.42	.30	-1.07	.83	-1.28
	C/T	.22	.42	.52	-.58	.83	-.69
	DC/C	.14	.42	.33	-.16	.83	-.19
	CT/T	.72	.42	1.71	.05	.83	.06
	CP/T	-.01	.42	-.02	-.92	.83	-1.10
	CN/T	.37	.42	.88	.12	.83	.14
ME	MLS	.21	.42	.5	-.54	.83	-.65
	MLT	.13	.42	.30	.14	.83	.16
	MLC	-.03	.42	-.07	-1.05	.83	-1.26
	C/S	-.24	.42	-.57	-.83	.83	-1
	VP/T	.52	.42	1.23	-.12	.83	-.14
	C/T	.03	.42	.07	-.55	.83	-.66
	DC/C	-.19	.42	-.45	.02	.83	.02
	CT/T	.30	.42	.71	-.31	.83	-.37
	CP/T	-.05	.42	-.11	-.19	.83	-.22
	CN/T	-.07	.42	-.16	-.75	.83	-.90
NF	MLS	.05	.42	.11	-1.13	.83	-1.36

MLT	.06	.42	.14	-.56	.83	-.67
MLC	.07	.42	.16	-.68	.83	-.81
C/S	-.50	.42	-1.19	.35	.83	.42
VP/T	-.30	.42	-.71	-.73	.83	-.87
C/T	-.42	.42	-1	-.48	.83	-.57
DC/C	-.38	.42	-.90	-.53	.83	-.63
CT/T	.25	.42	.59	-.17	.83	-.20
CP/T	.15	.42	.35	-.67	.83	-.80
CN/T	.30	.42	.71	-.87	.83	-1.04

Shapiro-Wilk Test of Normality for Pretest of Syntactic Complexity

Table 4.52

Shapiro-Wilk Test of Normality for Pretest of Syntactic Complexity

Indices / Group	MLS	MLT	MLC	C/S	VP/T	C/T	DC/C	CT/T	CP/T	CN/T
DCF	.18	.89	.72	.08	.30	.58	.80	.09	.46	.45
ME	.16	.99	.46	.14	.37	.57	.80	.37	.20	.71
NF	.07	.71	.37	.24	.34	.22	.55	.68	.77	.38

Appendix 11. Normality Tests for Posttest of Syntactic Complexity

Table 4.53

Normality Tests for Posttest of Syntactic Complexity

Group		Skewness			Kurtosis		
		Statistic	Std. Error	Normality	Statistic	Std. Error	Normality
DCF	MLS	.39	.42	.92	-.53	.83	-.63
	MLT	.33	.42	.78	-.58	.83	-.69
	MLC	.09	.42	.21	-.02	.83	-.02
	C/S	.15	.42	.35	-.09	.83	-.10
	VP/T	-.10	.42	-.23	-.99	.83	-1.19
	C/T	-.41	.42	-.97	-.11	.83	-.13
	DC/C	.40	.42	.95	-.04	.83	-.04
	CT/T	.35	.42	.83	-.19	.83	-.22
	CP/T	-.25	.42	-.59	.02	.83	.02
	CN/T	.55	.42	1.30	.50	.83	.60
ME	MLS	.04	.42	.09	-.85	.83	-1.02
	MLT	.02	.42	.04	-.26	.83	-.31
	MLC	-.17	.42	-.40	-.82	.83	-.98
	C/S	.01	.42	.02	.83	.83	1
	VP/T	-.02	.42	-.04	-.44	.83	-.53
	C/T	-.17	.42	-.40	-.76	.83	-.91
	DC/C	-.10	.42	-.23	-.12	.83	-.14
	CT/T	.40	.42	.95	-.28	.83	-.33
	CP/T	.18	.42	.42	-.12	.83	-.14
	CN/T	.18	.42	.42	.22	.83	.26
NF	MLS	.04	.42	.09	-.48	.83	-1.14

MLT	.08	.42	.19	-.45	.83	-.54
MLC	-.05	.42	-.11	-.90	.83	-1.08
C/S	.07	.42	.16	-.37	.83	-.44
VP/T	.36	.42	.85	-.43	.83	-.51
C/T	.34	.42	.80	-.63	.83	-.75
DC/C	.18	.42	.42	-.46	.83	-.55
CT/T	.19	.42	.45	-.48	.83	-.57
CP/T	-.02	.42	-.04	-.37	.83	-.44
CN/T	.04	.42	.09	-.92	.83	-1.10

Shapiro-Wilk Test of Normality for Posttest of Syntactic Complexity

Table 4.54

Shapiro-Wilk Test of Normality for Posttest of Syntactic Complexity

Indices / Group	MLS	MLT	MLC	C/S	VP/T	C/T	DC/C	CT/T	CP/T	CN/T
DCF	.36	.44	.94	.90	.28	.82	.66	.32	.60	.58
ME	.38	.44	.47	.46	.94	.23	.75	.56	.75	.49
NF	.66	.92	.55	.25	.70	.21	.89	.82	.95	.42

Normality Tests for Delayed-Posttest of Syntactic Complexity

Table 4.55

Normality Tests for Delayed-Posttest of Syntactic Complexity

Group		Skewness			Kurtosis		
		Statistic	Std. Error	Normality	Statistic	Std. Error	Normality
DCF	MLS	.43	.42	1.02	-.58	.83	-.69
	MLT	.32	.42	.76	-.58	.83	-.69
	MLC	.09	.42	.21	-.03	.83	-.03
	C/S	.01	.42	.02	-.29	.83	-.34
	VP/T	.34	.42	.80	-.47	.83	-.56
	C/T	-.40	.42	-.95	-.17	.83	-.20
	DC/C	.40	.42	.95	-.04	.83	-.04
	CT/T	.35	.42	.83	-.20	.83	-.24
	CP/T	.49	.42	1.16	-.45	.83	-.54
	CN/T	.55	.42	1.30	.50	.83	.60
ME	MLS	.05	.42	.11	-.86	.83	-1.03
	MLT	.04	.42	.09	-.21	.83	-.25
	MLC	-.17	.42	-.40	-.82	.83	-.98
	C/S	.02	.42	.04	.82	.83	.98
	VP/T	-.18	.42	-.42	-.74	.83	-.89
	C/T	-.17	.42	-.40	-.76	.83	-.91
	DC/C	-.10	.42	-.23	-.12	.83	-.14
	CT/T	.40	.42	.95	-.28	.83	-.33
	CP/T	.31	.42	.73	-.60	.83	-.72
	CN/T	.18	.42	.42	.22	.83	.26
NF	MLS	.04	.42	.09	-.47	.83	-.56

MLT	.07	.42	.16	-.47	.83	-.56
MLC	-.05	.42	-.11	-.90	.83	-1.08
C/S	.08	.42	.19	-.33	.83	-.39
VP/T	.13	.42	.30	-.70	.83	-.84
C/T	.34	.42	.80	-.62	.83	-.74
DC/C	.18	.42	.42	-.46	.83	-.55
CT/T	.19	.42	.45	-.48	.83	-.57
CP/T	-.28	.42	-.66	-.27	.83	-.32
CN/T	.04	.42	.09	-.91	.83	-1.09

Shapiro-Wilk Test of Normality for Delayed-Posttest of Syntactic Complexity

Table 4.56

Shapiro-Wilk Test of Normality for Delayed-Posttest of Syntactic Complexity

Indices / Group	MLS	MLT	MLC	C/S	VP/T	C/T	DC/C	CT/T	CP/T	CN/T
DCF	.23	.44	.94	.95	.76	.81	.66	.32	.27	.58
ME	.38	.34	.47	.49	.70	.23	.75	.56	.49	.49
NF	.69	.91	.55	.71	.66	.22	.89	.82	.86	.42

Appendix 12. Correlations between Dependent Variables for Pretest of Syntactic Complexity

Table 4.57

Correlations between Dependent Variables for Pretest of Syntactic Complexity

Correlations

		MLS	MLT	MLC	C/S	VP/T	C/T	DC/C	CT/T	CP/T	CN/T
MLS	Pearson Correlation	1	.779**	.153	.562**	.562**	.563**	.373**	.535**	.025	.486**
	Sig. (2-tailed)		.000	.149	.000	.000	.000	.000	.000	.813	.000
	N	90	90	90	90	90	90	90	90	90	90
MLT	Pearson Correlation	.779**	1	.223*	.332**	.600**	.478**	.311**	.504**	.214*	.571**
	Sig. (2-tailed)	.000		.034	.001	.000	.000	.003	.000	.043	.000
	N	90	90	90	90	90	90	90	90	90	90
MLC	Pearson Correlation	.153	.223*	1	-.456**	-.179	-.561**	-.326**	-.348**	.162	.313**
	Sig. (2-tailed)	.149	.034		.000	.091	.000	.002	.001	.128	.003
	N	90	90	90	90	90	90	90	90	90	90
C/S	Pearson Correlation	.562**	.332**	-.456**	1	.521**	.770**	.545**	.495**	-.094	.101
	Sig. (2-tailed)	.000	.001	.000		.000	.000	.000	.000	.377	.344
	N	90	90	90	90	90	90	90	90	90	90
VP/T	Pearson Correlation	.562**	.600**	-.179	.521**	1	.737**	.589**	.573**	.109	.307**
	Sig. (2-tailed)	.000	.000	.091	.000		.000	.000	.000	.307	.003
	N	90	90	90	90	90	90	90	90	90	90

C/T	Pearson Correlation	.563**	.478**	-.561**	.770**	.737**	1	.682**	.704**	-.022	.258*
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000	.000	.835	.014
	N	90	90	90	90	90	90	90	90	90	90
DC/C	Pearson Correlation	.373**	.311**	-.326**	.545**	.589**	.682**	1	.523**	.159	.154
	Sig. (2-tailed)	.000	.003	.002	.000	.000	.000		.000	.135	.147
	N	90	90	90	90	90	90	90	90	90	90
CT/T	Pearson Correlation	.535**	.504**	-.348**	.495**	.573**	.704**	.523**	1	.089	.337**
	Sig. (2-tailed)	.000	.000	.001	.000	.000	.000	.000		.402	.001
	N	90	90	90	90	90	90	90	90	90	90
CP/T	Pearson Correlation	.025	.214*	.162	-.094	.109	-.022	.159	.089	1	.159
	Sig. (2-tailed)	.813	.043	.128	.377	.307	.835	.135	.402		.135
	N	90	90	90	90	90	90	90	90	90	90
CN/T	Pearson Correlation	.486**	.571**	.313**	.101	.307**	.258*	.154	.337**	.159	1
	Sig. (2-tailed)	.000	.000	.003	.344	.003	.014	.147	.001	.135	
	N	90	90	90	90	90	90	90	90	90	90

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Appendix 13. Scatterplots for Linearity Assumption for Pretest of Syntactic Complexity

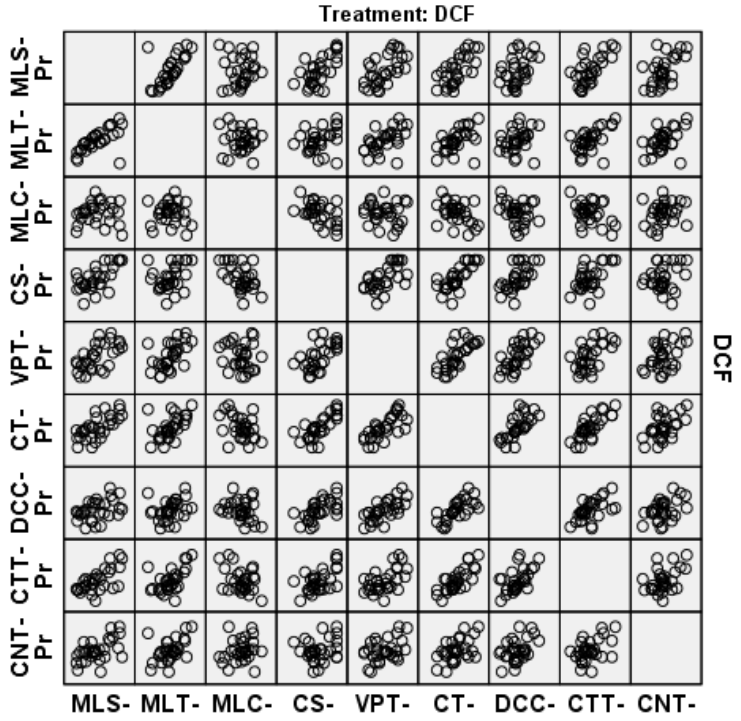


Figure 4.105. Linearity Assumption for Pretest of DCF

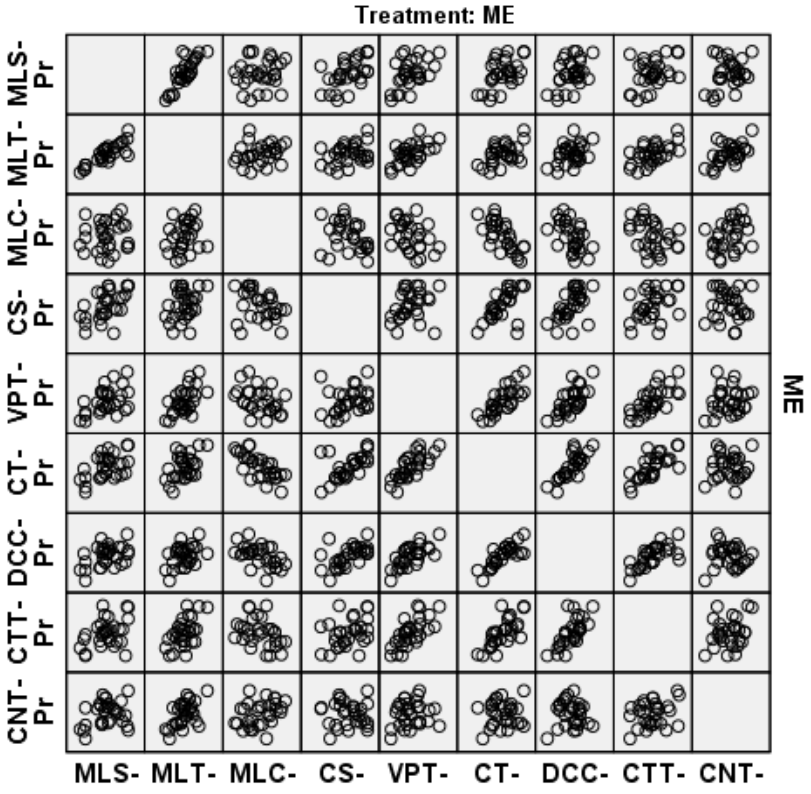


Figure 4.106. Linearity Assumption for Pretest of ME

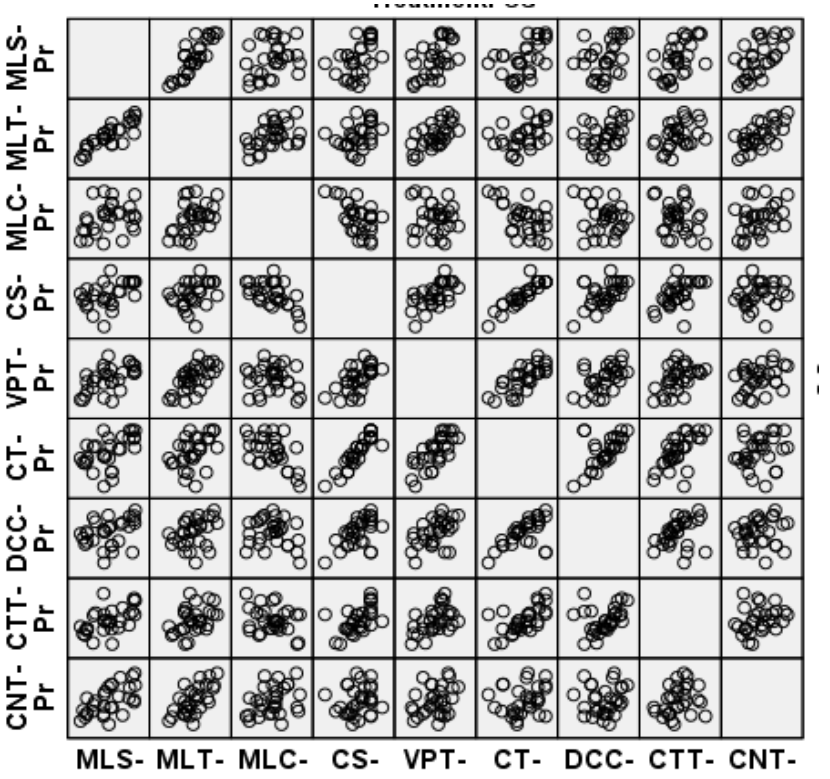


Figure 4.107. Linearity Assumption for Pretest of NF

Appendix 14. Correlations between Dependent Variables for Posttest of Syntactic Complexity

Table 4.58

Correlations between Dependent Variables for Posttest of Syntactic Complexity

Correlations

		MLS	MLT	MLC	C/S	VP/T	C/T	DC/C	CT/T	CP/T	CN/T
MLS	Pearson Correlation	1	.785**	.060	.222*	.612**	.329**	.487**	.558**	.022	.379**
	Sig. (2-tailed)		.000	.576	.035	.000	.002	.000	.000	.835	.000
	N	90	90	90	90	90	90	90	90	90	90
MLT	Pearson Correlation	.785**	1	.189	.117	.646**	.311**	.353**	.470**	.066	.491**
	Sig. (2-tailed)	.000		.074	.272	.000	.003	.001	.000	.540	.000
	N	90	90	90	90	90	90	90	90	90	90
MLC	Pearson Correlation	.060	.189	1	-.130	-.177	-.318**	-.541**	-.509**	-.081	.307**
	Sig. (2-tailed)	.576	.074		.224	.095	.002	.000	.000	.450	.003
	N	90	90	90	90	90	90	90	90	90	90
C/S	Pearson Correlation	.222*	.117	-.130	1	.254*	.675**	.179	.101	.170	.042
	Sig. (2-tailed)	.035	.272	.224		.016	.000	.092	.345	.109	.694
	N	90	90	90	90	90	90	90	90	90	90
VP/T	Pearson Correlation	.612**	.646**	-.177	.254*	1	.522**	.608**	.554**	.066	.231*
	Sig. (2-tailed)	.000	.000	.095	.016		.000	.000	.000	.534	.029
	N	90	90	90	90	90	90	90	90	90	90

C/T	Pearson Correlation	.329**	.311**	-.318**	.675**	.522**	1	.409**	.377**	.108	.258*
	Sig. (2-tailed)	.002	.003	.002	.000	.000		.000	.000	.310	.014
	N	90	90	90	90	90	90	90	90	90	90
DC/C	Pearson Correlation	.487**	.353**	-.541**	.179	.608**	.409**	1	.790**	.090	.189
	Sig. (2-tailed)	.000	.001	.000	.092	.000	.000		.000	.401	.074
	N	90	90	90	90	90	90	90	90	90	90
CT/T	Pearson Correlation	.558**	.470**	-.509**	.101	.554**	.377**	.790**	1	.035	.240*
	Sig. (2-tailed)	.000	.000	.000	.345	.000	.000	.000		.744	.023
	N	90	90	90	90	90	90	90	90	90	90
CP/T	Pearson Correlation	.022	.066	-.081	.170	.066	.108	.090	.035	1	-.014
	Sig. (2-tailed)	.835	.540	.450	.109	.534	.310	.401	.744		.892
	N	90	90	90	90	90	90	90	90	90	90
CN/T	Pearson Correlation	.379**	.491**	.307**	.042	.231*	.258*	.189	.240*	-.014	1
	Sig. (2-tailed)	.000	.000	.003	.694	.029	.014	.074	.023	.892	
	N	90	90	90	90	90	90	90	90	90	90

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Appendix 15. Scatterplots for Linearity Assumption for Posttest of Syntactic Complexity

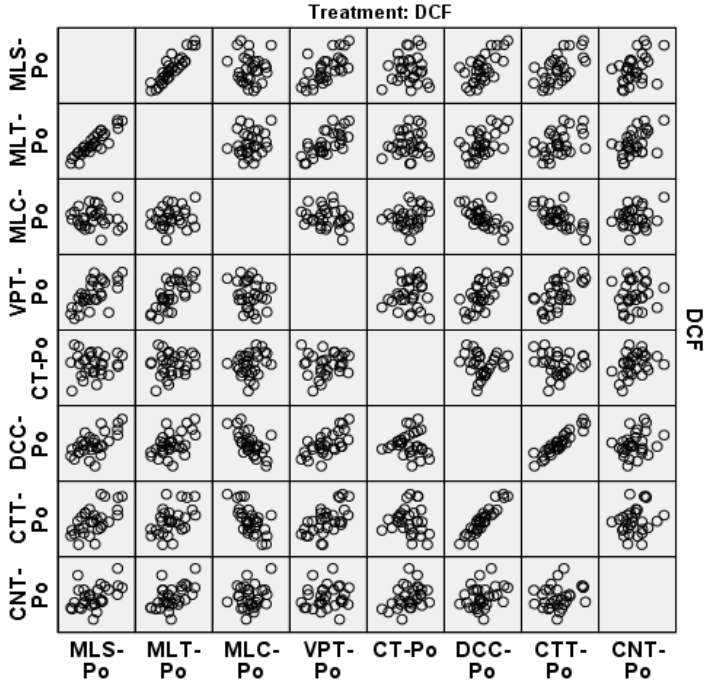


Figure 4.108. Linearity Assumption for Posttest of DCF

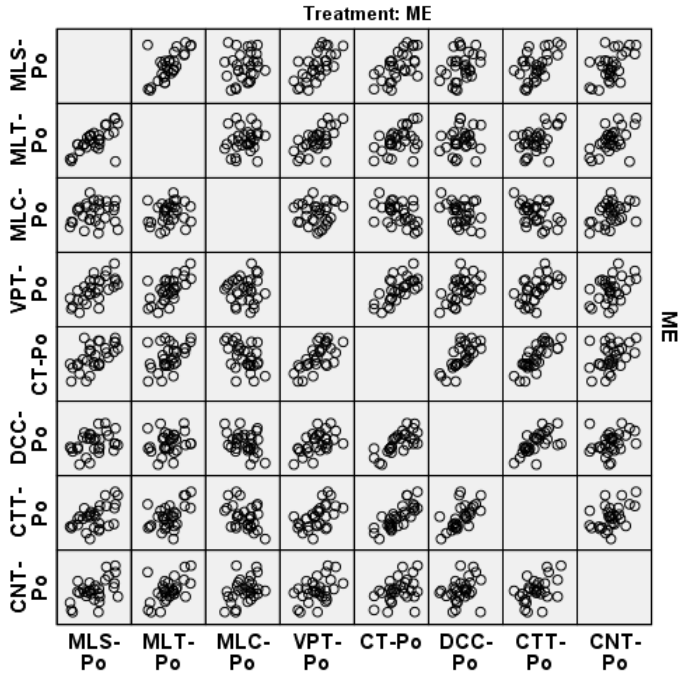


Figure 4.109. Linearity Assumption for Posttest of ME

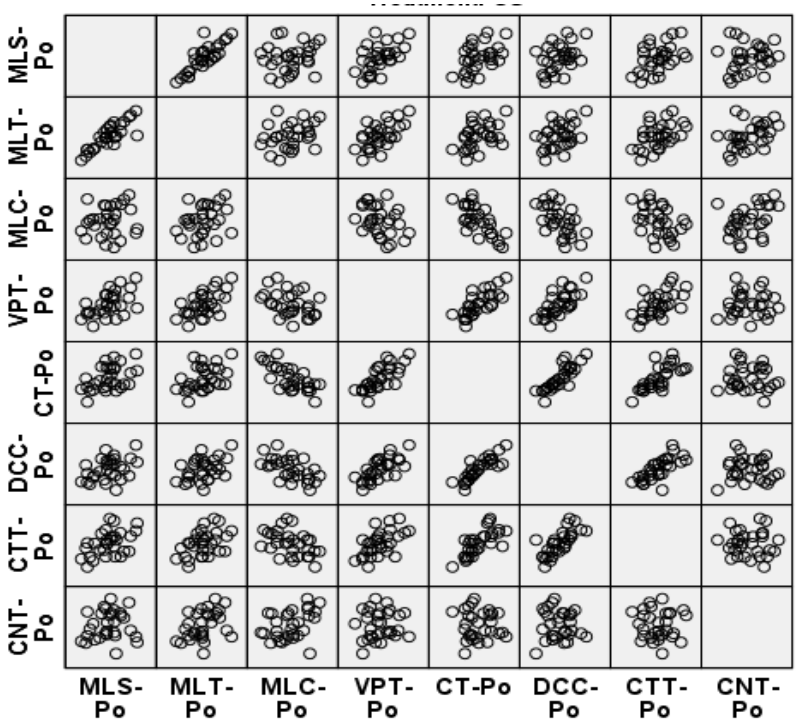


Figure 4.110. Linearity Assumption for Posttest of NF

Appendix 16. Correlations between Dependent Variables for Delayed-Posttest of Syntactic Complexity

Table 5.59

Correlations between Dependent Variables for Delayed-Posttest of Syntactic Complexity

Correlations

		MLS	MLT	MLC	C/S	VP/T	C/T	DC/C	CT/T	CP/T	CN/T
MLS	Pearson Correlation	1	.777**	.063	.226*	.418**	.328**	.485**	.556**	.207	.376**
	Sig. (2-tailed)		.000	.555	.032	.000	.002	.000	.000	.051	.000
	N	90	90	90	90	90	90	90	90	90	90
MLT	Pearson Correlation	.777**	1	.189	.118	.369**	.309**	.356**	.474**	.329**	.492**
	Sig. (2-tailed)	.000		.074	.269	.000	.003	.001	.000	.002	.000
	N	90	90	90	90	90	90	90	90	90	90
MLC	Pearson Correlation	.063	.189	1	-.135	-.124	-.319**	-.541**	-.508**	.002	.307**
	Sig. (2-tailed)	.555	.074		.204	.245	.002	.000	.000	.987	.003
	N	90	90	90	90	90	90	90	90	90	90
C/S	Pearson Correlation	.226*	.118	-.135	1	.118	.686**	.171	.102	.151	.059
	Sig. (2-tailed)	.032	.269	.204		.266	.000	.108	.338	.155	.582
	N	90	90	90	90	90	90	90	90	90	90
VP/T	Pearson Correlation	.418**	.369**	-.124	.118	1	.313**	.438**	.445**	-.026	.206
	Sig. (2-tailed)	.000	.000	.245	.266		.003	.000	.000	.810	.051
	N	90	90	90	90	90	90	90	90	90	90
C/T	Pearson Correlation	.328**	.309**	-.319**	.686**	.313**	1	.409**	.378**	.140	.257*

	Sig. (2-tailed)	.002	.003	.002	.000	.003		.000	.000	.188	.015
	N	90	90	90	90	90	90	90	90	90	90
DC/C	Pearson Correlation	.485**	.356**	-.541**	.171	.438**	.409**	1	.790**	.283**	.189
	Sig. (2-tailed)	.000	.001	.000	.108	.000	.000		.000	.007	.074
	N	90	90	90	90	90	90	90	90	90	90
CT/T	Pearson Correlation	.556**	.474**	-.508**	.102	.445**	.378**	.790**	1	.195	.240*
	Sig. (2-tailed)	.000	.000	.000	.338	.000	.000	.000		.066	.023
	N	90	90	90	90	90	90	90	90	90	90
CP/T	Pearson Correlation	.207	.329**	.002	.151	-.026	.140	.283**	.195	1	.183
	Sig. (2-tailed)	.051	.002	.987	.155	.810	.188	.007	.066		.084
	N	90	90	90	90	90	90	90	90	90	90
CN/T	Pearson Correlation	.376**	.492**	.307**	.059	.206	.257*	.189	.240*	.183	1
	Sig. (2-tailed)	.000	.000	.003	.582	.051	.015	.074	.023	.084	
	N	90	90	90	90	90	90	90	90	90	90

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Appendix 17. Linearity Assumption for Delayed-Posttest of Syntactic Complexity

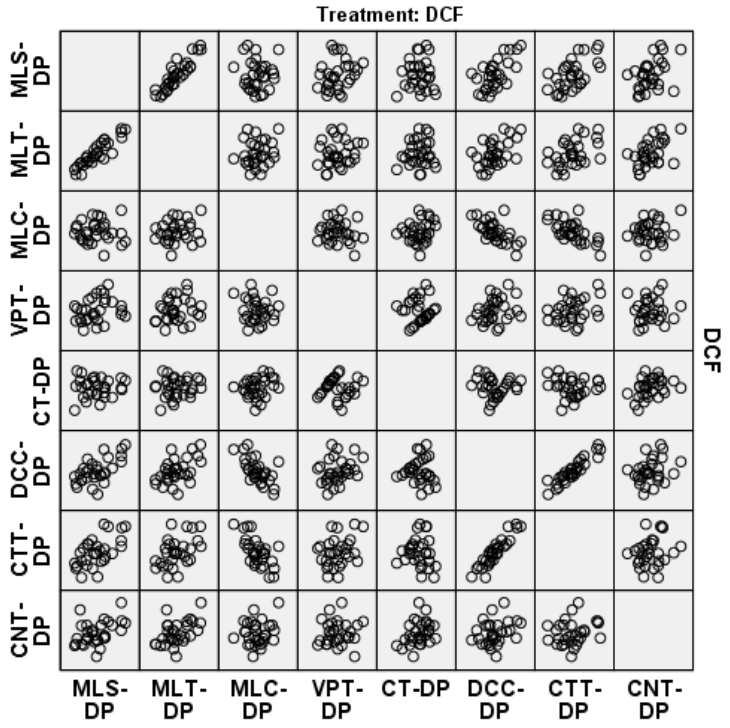


Figure 4.111. Linearity Assumption for Delayed-Posttest of DCF

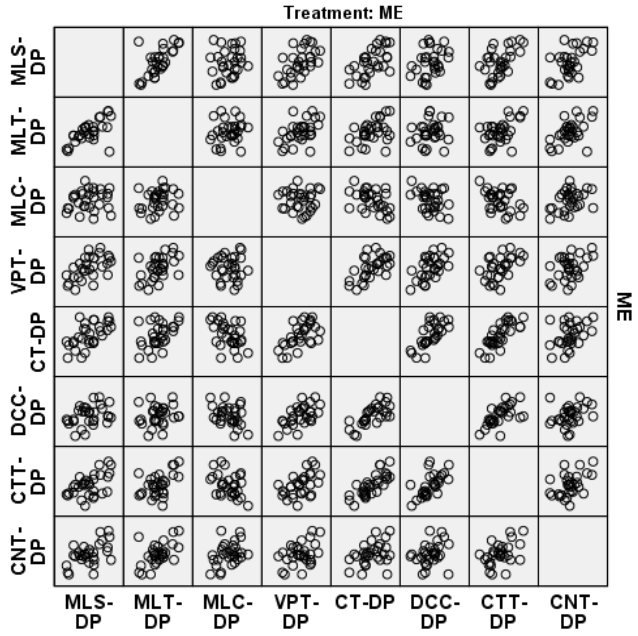


Figure 4.112. Linearity Assumption for Delayed-Posttest of ME

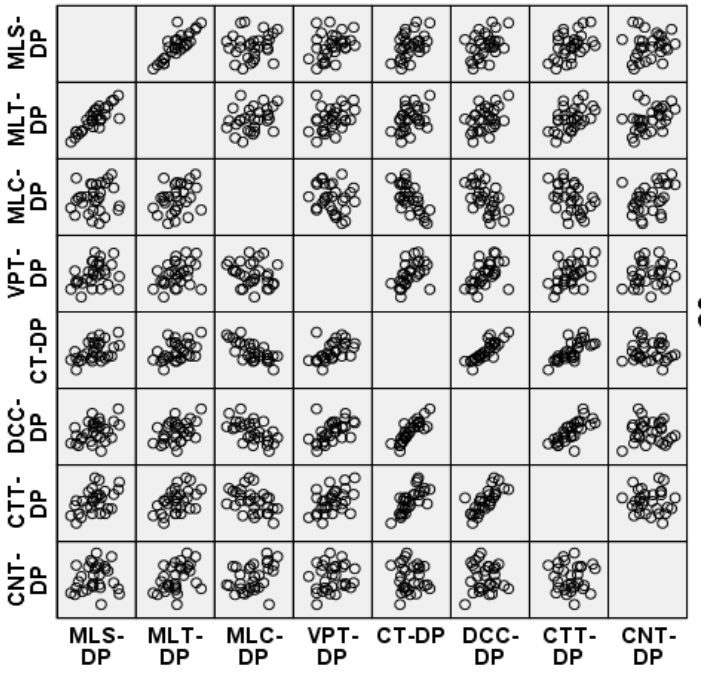


Figure 4.113. Linearity Assumption for Delayed-Posttest of NF



GAZİLİ OLMAK AYRICALIKTIR..