**T.C. KOCAELİ ÜNİVERSİTESİ**
**SOSYAL BİLİMLER ENSTİTÜSÜ**
**YABANCI DİLLER EĞİTİMİ ANABİLİM DALI**
**İNGİLİZ DİLİ EĞİTİMİ BİLİM DALI**

# COMPUTERIZED DYNAMIC ASSESSMENT OF SECOND LANGUAGE READING COMPREHENSION: AN INVESTIGATION OF THE PREDICTIVE EFFECT OF WORKING MEMORY

**(YÜKSEK LİSANS TEZİ)**

**Merve BAYRAM**

**KOCAELİ 2021**

**T.C. KOCAELİ ÜNİVERSİTESİ**
**SOSYAL BİLİMLER ENSTİTÜSÜ**
**YABANCI DİLLER EĞİTİMİ ANABİLİM DALI**
**İNGİLİZ DİLİ EĞİTİMİ BİLİM DALI**

# COMPUTERIZED DYNAMIC ASSESSMENT OF SECOND LANGUAGE READING COMPREHENSION: AN INVESTIGATION OF THE PREDICTIVE EFFECT OF WORKING MEMORY

**(YÜKSEK LİSANS TEZİ)**

**Merve BAYRAM**

**Dr. Öğr. Üyesi Mehmet ALTAY**

**KOCAELİ 2021**

**T.C. KOCAELİ ÜNİVERSİTESİ**
**SOSYAL BİLİMLER ENSTİTÜSÜ**
**YABANCI DİLLER EĞİTİMİ ANABİLİM DALI**
**İNGİLİZ DİLİ EĞİTİMİ BİLİM DALI**

# COMPUTERIZED DYNAMIC ASSESSMENT OF SECOND LANGUAGE READING COMPREHENSION: AN INVESTIGATION OF THE PREDICTIVE EFFECT OF WORKING MEMORY

**(YÜKSEK LİSANS TEZİ)**

**Tezi Hazırlayan: Merve BAYRAM**

**Tezin Kabul Edildiği Enstitü Yönetim Kurulu Karar ve No: 23.06.2021/15**

**KOCAELİ 2021**

To my eyes, where I caused the myopic degrees to increase, and to my family, who trusts me more than I do…

## ACKNOWLEDGEMENTS

I would like to express my first and greatest gratitude to my advisor, Asst. Prof. Dr. Mehmet Altay for his valuable guidance not only in this thesis process but also in the rest of my master's degree. His continuous support, precious feedback, and advice at every stage have always encouraged me to research and learn more throughout this process. There are also other things that I have learned from him beyond his knowledge and experience. He supported me at such important points in my life that I always felt his support and belief in me. I am sincerely honored to be his student.

Secondly, I would like to show my appreciation to my committee members, Assoc. Prof. Dr. Doğan Yüksel and Asst. Prof Dr. Sezen Arslan for their valuable feedback and recommendations for this research. I also had the chance to be their student and I am grateful for everything I have learned from them. I want to extend my gratitude to Assoc. Prof. Dr. Doğan Yüksel for his guidance and contributions to my academic development. He also set an example for me to be always an understanding and helpful teacher. Moreover, he and Asst. Prof. Dr. Mehmet Altay gave me the opportunity to meet with dynamic assessment; therefore, I also appreciate them for it.

Although I cannot express their great effort in a few sentences, I would like to express my deepest gratitude to my parents, Rabiye and Necdet Bayram; my sisters, Ayşenur and Tuğba Bayram. Without their never-ending love and support, I would not be who and where I am right now. They encouraged me to go to the graduate exam of Kocaeli University at a time when I lost my hope and faith, and I was accepted into the program. Now, I am glad that you did not let my failures make me give up. Thank you all for trusting me more than I did!

I would like to thank the participants of the current research for their time and contribution. I also express my special thanks to my friends Hazal Dilan Fidan and Şifanur Demir for their time and feedback on the suitability of the passages and test items. I would like to extend my thanks to Şifanur Demir for being my best and most compatible partner in graduate school. In addition, very sincere gratitude goes to my friend Hatice Beyza Erdinç for always supporting and motivating me in both my best and worst times. Finally, I am grateful to all my previous teachers, from those in

primary school to those in graduate school, who have contributed to my academic success in some way.

## TABLE OF CONTENTS

## ÖZET

**Bilgisayarlı dinamik değerlendirme, Vygotsky'nin sosyokültürel kuramına dayanan dinamik değerlendirmenin pratiklik sorunlarına alternatif olarak geliştirilen, nispeten yeni bir yaklaşımdır. Dinamik değerlendirme, arabuluculuk yoluyla öğrencileri destekleyerek, değerlendirme ve öğretimi bütünleştirir. Önceki araştırmalar, ikinci/yabancı dil gelişimini teşhis ve teşvik etmedeki rolünü göstermiştir. Öte yandan, çalışma belleği, ikinci dil işleme sırasında bilgilerin depolanması ve işlenmesinde önemli bir rol oynar ve ikinci dil okuma becerisi ile ilişkilendirilmiştir. Bu durum ve bilgisayarlı dinamik değerlendirme ile ilgili sınırlı araştırmadan esinlenen bu araştırma, öncelikle çalışma belleğinin müdahaleci bir bilgisayarlı dinamik değerlendirmede gerçek ve aracılı ikinci dil okuduğunu anlama puanları üzerindeki yordayıcı etkisini araştırmayı amaçlamaktadır. Diğer bir amaç ise, öğrencilerin gerçek ve aracılı puanlarını karşılaştırarak, bilgisayarlı dinamik değerlendirme prosedürünün okuma performansı üzerindeki etkisini analiz etmektir. Ayrıca, arabuluculuk kullanım düzeyi ve öğrenme potansiyeli puanının öğrenciler arasında nasıl farklılaştığını incelemeyi amaçlamaktadır. Veriler, bilgisayarlı bir otomatik okuma aralığı görevi (RSPAN) ve araştırmacı tarafından geliştirilen ikinci dilde okuduğunu anlamanın bilgisayarlı dinamik değerlendirmesi ile nicel olarak toplanmıştır. Araştırmaya, İngilizce Öğretmenliği Bölümü'nde akademik amaçlı İngilizce öğrenen 59 öğrenci katılmıştır. Sonuçlar, SPSS 18.0 ile yapılan veri analizine dayanarak, çalışma belleğinin gerçek veya aracılı ikinci dil okuduğunu anlama puanlarını önemli ölçüde öngörmediğini ortaya koydu; bununla birlikte bilgisayarlı dinamik değerlendirme prosedürü, çalışma belleğinin ikinci dil okumaya katılımını biraz arttırdı. Diğer taraftan, öğrenciler, arabuluculuk kullanımları ve öğrenme potansiyeli puanları farklı olmasına rağmen, arabuluculuk sürecinden önemli ölçüde yararlanmışlardır. Bulgular ışığında, çeşitli pedagojik çıkarımlar da tartışılmıştır.**



**Anahtar Kelimeler:** Bilgisayarlı dinamik değerlendirme, dinamik değerlendirme, sosyokültürel kuram, arabuluculuk, çalışma belleği, ikinci dil okuduğunu anlama

## ABSTRACT

**Computerized dynamic assessment (C-DA) is a relatively new approach developed as an alternative to the practicality issues of dynamic assessment (DA), which is rooted in Vygotsky's sociocultural theory. DA integrates assessment and instruction by supporting learners through mediation. Previous research indicated its role in diagnosing and promoting second/foreign language (L2) development. On the other side, working memory (WM) plays an important role in storing and processing information during L2 processing and has been associated with L2 reading ability. Inspired by this fact and limited research on C-DA, this research primarily aims to investigate the predictive effect of WM on actual and mediated L2 reading comprehension scores in an interventionist C-DA. Another aim is to analyze the effect of the C-DA procedure on students' reading performance by comparing the actual and mediated scores. Furthermore, it seeks to examine the level of mediation use and how learning potential score (LPS) differs among the students. The data were collected quantitatively through a computerized automated reading span task (RSPAN) and a C-DA of L2 reading comprehension developed by the researcher. Fifty-nine students learning English for academic purposes in English Language Teaching Department participated in the research. Based on the data analysis through SPSS 18.0, the results revealed that WM did not significantly predict either actual or mediated L2 reading comprehension scores; however, the C-DA procedure slightly increased the involvement of WM in L2 reading. On the other hand, students significantly benefited from the mediation process although their mediation use and LPSs varied. In light of the findings, several pedagogical implications were also discussed.**



**Keywords:** Computerized dynamic assessment, dynamic assessment, sociocultural theory, mediation, working memory, second lanugage reading comprehension

## LIST OF ABBREVIATIONS

**DA: Dynamic Assessment**

**C-DA: Computerized Dynamic Assessment**

**EAP: English for Academic Purposes**

**EFL: English as a Foreign Language**

**ELT: English Language Teaching**

**ESL: English as a second language**

**GRE: Graduate Record Examination**

**L1: First Language**

**L2: Second Language**

**LPS: Learning Potential Score**

**MLE: Mediated Learning Experience**

**N-DA: Non-Dynamic Assessment**

**RST: Reading Span Task**

**RSPAN: Computerized Automated Reading Span Task**

**SCT: Socio-cultural Theory**

**STM: Short-term Memory**

**SPSS: Statistical Package for the Social Sciences**

**WM: Working Memory**

**ZAD: Zone of Actual Development**

**ZPD: Zone of Proximal Development**

## LIST OF TABLES

## LIST OF FIGURES

# INTRODUCTION

Dynamic Assessment (DA) has its origins in Vygotsky's sociocultural theory (SCT), especially his idea of Zone of Proximal Development (ZPD), unlike the traditional static-based assessments that are based on psychometric principles (Poehner & Lantolf, 2005: p. 235). DA is built on the idea that human abilities arise in the activities in which learners are guided by other people and through the use of cultural tools (Poehner & Lantolf, 2013: p. 324). It includes a procedure where learners are given any kind of support, which is mediation, during the assessment when they have difficulty (Haywood & Lidz, 2007: p. 13). Learners' responsiveness to mediation informs not only their current/ developed abilities but also their future performance. Moreover, it helps them to develop their ZPD which "is not what has developed but what is developing" (Grigorenko & Sternberg, 1998: p. 78). Regarding the ZPD that occurs in the interaction, DA is a type of assessment that combines the process of assessment and instruction to determine learners' future performance. With the development of technology, researchers have recently begun to implement the principles of SCT in second or foreign language research through computers (e.g., Lee, 2008: p. 53; Li & Zhu, 2013: p. 61; Ma, 2017: p. 183). Moreover, Poehner (2008: p. 177) recommended computerized dynamic assessment (C-DA) as an alternative and a solution to the concerns of the ordinary DA procedures. DA and C-DA are based on integrating the mediational prompts to the procedure to gain insights about learners' developmental levels and to examine its effect on promoting improvement in various skills including second language (L2) comprehension and production.

On the other hand, the studies on language comprehension reported the variations in especially learners' reading comprehension. Daneman and Carpenter (1980: p. 450) stated that individual differences in language comprehension may be associated with learners' different working memory (WM) capacity, which is a dynamic system responsible for storing information for a limited time while dealing with other cognitive tasks (Savage et al., 2006: p. 186). Therefore, this research was designed to investigate whether students' WM scores are predictors of their non-dynamic (actual) and dynamic (mediated) reading comprehension scores obtained

from a C-DA. It also aims to compare their actual and mediated scores as well as their level of mediation use in the C-DA of reading comprehension. Finally, it also seeks to provide information about students' potential for learning through calculating Learning Potential Score (LPS). The data were gathered quantitatively through a computerized automated reading span task (RSPAN) as a WM measure and a C-DA of L2 reading comprehension. Overall analyses of the results indicated that WM scores did not significantly predict either actual or mediated reading comprehension scores. On the other side, students significantly benefited from the C-DA procedure according to the findings.

# CHAPTER 1

## 1. INTRODUCTION

This section aims to address the brief background to the research. It then goes on to the problem statement, aim, and significance of the research. Further, operational definitions are also presented.

### 1.1. BACKGROUND TO THE RESEARCH

Language assessment is important in understanding learners' knowledge or ability and informing the strengths and weaknesses of the teaching method. However, traditional static assessments mostly focus on the product rather than the process of learning. In the static assessments, learners are asked questions and rated based on their independent performance, namely without any interference (Tzuriel, 2001: p. 13), thus assessment and instruction are separated. As a solution to these problems, DA is developed to assess "an individual's potential for future development by embedding instruction in the assessment process itself" (Lantolf & Poehner, 2010: p. 13). That is, DA includes the process in which assessors actively provide intervention and evaluate the learner responsiveness to this intervention (Haywood & Lidz, 2007: p. 26). Therefore, DA enables the integration of assessment and instruction.

DA developed out of Vygotsky's SCT and especially his notion of ZPD which is described as the difference between what an individual can do with and without the guidance of a more knowledgeable person. According to Vygotsky, learning is viewed as an activity that is socially constructed (Leont'ev, 1981: pp. 37-71) and development occurs as individuals interact and collaborate with others in social activities, which enables them to internalize these interactions. What a child can do

with other people's assistance helps us get information about his mental development more than what he can do alone, and ZPD allows us the information about a child's future and dynamic development (Vygotsky, 1978: p. 131). In this regard, DA is based on the idea that the abilities emerge from the activities where learners are supported by other people as well as cultural tools (Poehner & Lantolf, 2013: p. 324). It is also dialogic in its nature because it includes a procedure that learners are provided mediation, which is any kind of assistance, in accordance with learners' ZPD when they encounter difficulty during the assessment. From a Vygotskian perspective, DA reveals not only the developed abilities of learners but also their developing functions (Lantolf & Poehner 2004: p. 54). Because ZPD focuses on the abilities that are still in the formation process, it also provides important information on teaching (Vygotsky, 1987, as cited in Yang & Qian, 2019: p. 3). In other words, DA aims to improve students' ZPD through the mediation process and encourage further development. Needless to say, mediational prompts are determined based on learners' ZPD, thus there are some practicality concerns about DA procedure. Therefore, Poehner (2008: p. 177) recommended the C-DA which offers computer-integrated mediation during the assessment on various language skills such as reading comprehension.

DA is an effective way of revealing and promoting development. In particular, this groundbreaking type of assessment is associated with a new understanding of the instruction and assessment in reading ability (Pishghadam & Barabadi, 2012: p. 74). For instance, Afflerbach (2007: p. 32) claims that the continuation of reading development should be the key outcome of while assessing reading. During reading, individuals benefit from various cognitive processes such as storing or retrieving information and skills or strategy use, which can cause variations in the reading performances of learners. Further, individual differences in reading comprehension are thought to be related to the differences in the storage and processing functions of WM capacity (Daneman & Carpenter, 1980: p. 450). In this respect, many researchers (e.g., Alptekin & Erçetin, 2009: p. 627, 2010: p. 206; Harrington & Sawyer, 1992: p. 25; Jeon &Yamashita, 2014: p. 160; Linck et al., 2014: p. 861; Walter, 2004: p. 315) reported the positive relationship of L2 reading comprehension

with WM, which is “an integrated system for temporarily storing and manipulating information” (Baddeley, 2003: p. 837). In light of the issues mentioned, the current thesis focuses on the predictive role of WM in learners’ actual and mediated scores in a C-DA of reading comprehension in L2. It also aims to investigate the impact of the C-DA procedure on students' reading performance, the use of mediation among the students, and their learning potential.

## 1.2. STATEMENT OF THE PROBLEM

The main problem of this thesis emerges from the inefficacy and inadequacy of static assessments, namely non-dynamic assessments (N-DA). Luria (1961: p. 7), in his pioneering study, compared the statistical and dynamic assessment procedures emphasizing the important factor of assistance during the assessment. Statistical assessments focus on the independent performance of learners representing the current level of development. As in many countries, students in Turkey are mostly assessed without any intervention, even receiving no feedback about their mistakes or lack of knowledge. Especially low achievers are discouraged by their scores, which increases their anxiety during examinations. Moreover, they do not receive any assistance to improve their performance and catch up with high achievers. Even the students with the same independent scores may differ in their learning potential and therefore instruction should also be based on their different needs. In this regard, Vygotsky (1978: pp. 85-86) stated that learners’ only Zone of Actual Development (ZAD) determined by their independent performance can not explain the whole picture of their development, thus he explained ZPD since their potential ability exceeds their actual ability when more knowledgeable others support learning (Wertsch, 1993: p. 28). DA also in line with the idea that a learner’s independent performance does not predict her future performance; that is, potential development (Poehner et al., 2017: p. 245).

Especially in language learning, the teacher is the only source of L2 input because students learn English as a Foreign Language (EFL), therefore getting sufficient feedback and assistance on their mistakes are crucial to improve their

abilities. DA allows language teachers to examine students' performances depending on their responsiveness to mediation. Thus, the mutual interaction between assessors and students combines teaching and assessment (Poehner, 2009: pp. 471-472). Taken together, DA provides a solution to learn about both actual and potential levels of development and it not only enhances students' performance in the given task but also promotes future development by internalizing the knowledge. The concerns about the applicability of DA motivate some researchers to implement C-DA, which is more practical in terms of workload and time, and offers comparable results for large groups of students.

On the other side, another concern behind the rationale of this thesis is neglecting the effect of WM on students' overall achievement, especially in L2 reading comprehension. During language processing, WM plays a crucial role in storing, manipulating, or integrating information. Especially some components of WM (i.e. the phonological loop and the central executive function) have been associated with some important aspects of language processing such as vocabulary acquisition and making inference (Baddeley, 2003: pp. 189-208). On this basis, many studies and meta-analyses highlight the meaningful relationship between WM and L2 reading ability (e.g., Daneman & Merikle, 1996: p. 422; Linck et al., 2014: p. 861; Shin, 2020: p. 873). However, on the one hand, studies with contradictory findings on this relationship have also raised questions about this issue, leading to the call for a further investigation.

## 1.3. AIM OF THE RESEARCH

DA has been a growing body of research in L2 learning and teaching although it is a relatively new field in Applied Linguistics (Poehner, 2008: p. 91). Similar to DA, the advantages of C-DA over traditional static assessment have been reported to date in L2 research. However, its applicability in educational settings and its impact on learners' performances and language development should be further investigated since the amount of research studying DA and C-DA in L2 contexts is quite limited in the Turkish context (e.g., Güleryüz-Adamhasan, 2019: pp. 124-129; Çalış, 2018:

pp. 89-96; Kır, 2020: pp. 103-105; Şentürk, 2019: pp. 72-74; Ulu, 2020: pp. 98-99; Yılmaz-Yakışık, 2012: pp. 129-136). Inspired by the information briefly mentioned and considering lack of research on C-DA as well as inconsistent findings on the role of WM in L2 reading comprehension, this thesis is based on two-way purpose related to L2 reading comprehension of the students currently learning English for academic purposes (EAP) in English Language Teaching Department (ELT).

The main goal of the current research is to determine whether students' WM performances predict their actual (independent) and mediated (with mediation) scores in a C-DA of L2 reading comprehension. Second, this thesis also set out to report the findings of the interventionist online C-DA of L2 reading comprehension and to reveal the effect of C-DA on students' performance. Regarding the second purpose, it primarily aims to investigate whether there is a significant difference between actual and mediated scores to measure the effect of mediational hints on their performance. Further, it seeks to examine whether the degree of mediation use among the students differs significantly during the C-DA. Finally, it also attempts to provide insights into the learning potential of students in L2 reading comprehension through the Learning Potential Score (LPS). Accordingly, how LPS differs among the students with the same actual score is also discussed.

## 1.4. SIGNIFICANCE OF THE RESEARCH

In educational psychology, DA is an area that is promising but still developing (Grigorenko, 2008: p.127). Therefore, the research areas of this thesis make important contributions to the literature in several ways. To begin with, the number of research investigating DA in L2 pedagogy is very limited in the Turkish context, thus there is a general lack of research in this regard. Moreover, the C-DA studies especially related to L2 reading ability are also relatively scarce. Therefore, there is a need for research to gain better insights into the applicability and effect of C-DA on learners' L2 reading comprehension. To the best knowledge of the researcher, no studies have been found regarding the implementation of C-DA to L2 reading in the Turkish context.

On the other hand, there is a vast array of research focusing on the relationship of WM capacity with L2 reading comprehension. There are also some researchers who have benefited from both WM and DA measures in their study (e.g., Stevenson et al., 2014: p. 51) or measured WM dynamically (e.g., Swanson, 2006: p. 125). However, no previous study has been conducted to date on the relationship between WM and DA or C-DA, i.e. the effect of WM on mediated scores of students determined by their responsiveness to mediation, hence this is also pioneering research in this sense. Finally, most of the DA studies in Turkey were conducted in EFL contexts (e.g., Kır, 2020: pp. 33-35; Ulu, 2020: p. 39); nevertheless, this research sheds light on the actual and potential reading abilities of the students learning EAP in the department of ELT.

## 1.5. OPERATIONAL DEFINITIONS

**Computerized Dynamic Assessment:** C-DA is an alternative type of DA that can be administered to large groups and reports of learners' performances automatically (Poehner, 2008: p. 177). C-DA was developed as a solution to ordinary DA with some practicality concerns and it is based on the idea that learners are provided pre-scripted mediational hints through computers when they respond incorrectly during the assessment, thus it mostly follows the interventionist DA approach.

**Dynamic Assessment:** DA is a new type of assessment where learners are provided spontaneous or pre-scripted mediation to reveal and promote their abilities, which integrates assessment and instruction (Poehner, 2008: p. 21).

**Mediation:** Mediation is any form of sequential assistance given in the form of an increasing series of explicitness when learners have difficulty during the assessment (Poehner, 2008: p. 38). During DA, learners can perform beyond their capabilities with the help of mediation, which creates instruction and promotes development (Poehner, 2009: p. 472).

**Non-Dynamic Assessment:** In N-DA, learners are evaluated without any intervention or assistance by the examiners (Sternberg & Grigorenko, 2002: p. 45)

and they only receive feedback about the score determined by their independent performance.

**Socio-cultural Theory:** According to Vygotsky (1978: pp. 24-30), knowledge is social and it is created in an atmosphere where learners cooperate and interact with other people. In SCT, human mental activities are mediated with the help of symbolic tools or social artifacts (Wells, 1999: pp. 60-71). Hence, learners encounter new knowledge through dialogues with others, build new understandings, and internalize them.

**Zone of Proximal Development**: Vygotsky (1978: p. 86) defined ZPD as "the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem-solving under guidance or in collaboration with more capable peers".

**Working Memory:** WM, a system with limited capacity, is responsible for the storage of temporary information during the activities that are cognitively demanding (Baddeley & Hitch, 1974: pp. 47-48).

# CHAPTER 2

## 2. LITERATURE REVIEW

The purpose of this chapter is to review the literature. First, an overview of SCT and its constructs is explained because DA is rooted in SCT. Second, a historical overview, definition, characteristics, and models of DA and DA studies in L2 contexts are elaborated. Finally, previous reseach on WM and its role in L2 reading comprehension are also discussed.

### 2.1. SOCIOCULTURAL THEORY

SCT, which is also known as "social historical theory, cultural psychology, and cultural historical psychology", was developed by Vygotsky and his colleagues in 1920s and 1930s (Poehner, 2008: p. 25). The emphasis in the theory is on social environment and cognition. SCT mainly assumes that any action is mediated in a social environment and it cannot be considered separately from this environment (Wertsch, 1993: p. 18). According to Vygotsky (1978: pp. 86-87), knowledge is social and it is constructed in an environment where learners collaborate, interact, and communicate with others. Vygotsky was inspired by Marxist philosophy, claimed that concrete activity in the environment mediated by physical tools shapes humans and Vygotsky and his colleagues proposed that human cognitive functions are also mediated, by extending this philosophy to the psychological plane (Leont'ev, 1981, as cited in Poehner, 2008: p. 25). With the help of interaction with other people and cultural objects, human cognition is mediated socially and culturally (Vygotsky, 1986, as cited in Bodrova & Leong, 2003: p. 160). That is, social interaction with more knowledgeable people and cultural tools generates changes in children's behavior and thought gradually.

From a Vygotskyan perspective, like other higher mental functions, learning is seen as a social activity and language learning is an activity that is also socially

constructed (Leont'ev, 1981: pp. 37-71). Moreover, Vygotsky believes that learning occurs when learners internalize social interactions and "there is a movement from the interpsychological plane (between individuals) to the intrapsychological plane (within an individual)" (McCarthy & McMahon, 1992: p. 18). This leads us to Vygotsky's well-known saying: "Every function in the child's cultural development appears twice: first, on the social level, and later, on the individual level; first, between people (interpsychological), and then inside the child (intrapsychological)" (Vygotsky, 1978: p. 57). He also stated that mental functions such as attention and logical memory are firstly between individuals as an "intermental" activity; then it becomes "intramental" activity for the individual (Vygotsky, 1981: p. 163).

Like Vygotsky, Halliday (1978: p. 24) claimed that language and social systems are interwoven. Through dialogues with others, learners face new knowledge, develop new understandings, and internalize them. Moreover, children improve their thinking processes in collaborative activities (Vygotsky, 1978: pp. 85-86). In SCT, human mental activities are mediated by symbolic tools or socially constructed artifacts, and language is one of the most important of them because linguistic abilities are important for some mental operations such as memory and attention (Wells, 1999: pp. 51- 97 ). In classroom discourse, language teachers take this role and assist their students with the help of collaborative interactions and activities. Thus, teachers help students reach a linguistic and conceptual understanding. Wells (1999: p. 137) also emphasized Vygotsky's theory of "collaborative community" in which students participate in the activities jointly. From the same perspective, Cazden (2001: p. 54) stated that a language classroom is both an academic and a social context.

### 2.1.1. Tools and Mediation

In the process of higher psychological development, people use various tools such as physical, symbolic, and psychological tools in their social and cultural environment (Kozulin, 2003: pp. 15-38). Physical tools created by humans can be found in different areas. People use these concrete tools to mediate their connection with the world; for example, we use some technological devices in educational

environments or we have eyeglasses, hearing devices, etc. in medicine (Poehner, 2008: p. 26).

On the other hand, we use some symbolic tools that Vygotsky called cultural artifacts. Lantolf (2000: p. 1) stated that "we use symbolic tools or signs to mediate and regulate our relationship with others and with ourselves and thus change the nature of these relationship". We use both simple tools such as pen, spoon, table, and more complex ones (i.e. cultural tools or artifacts) such as language, traditions, beliefs, and arts (Cole, 1997: pp. 144-145). Signs, symbols, formulas, graphic organizers, and texts can be given as other examples of symbolic tools or mediators that are more abstract. Hence, most of the cultural tools are language-based in Vygotsky's view (Bodrova & Leong, 2003: pp. 156-176). Symbolic tools, in sociocultural activities, are also highlighted by Vygotsky, with emphasis on formal education among these activities. According to Vygotsky, children's cognitive development and learning are attached to their mastery of symbolic mediators and internalization of them as psychological tools (Kozulin, 1998: p. 57). Moreover, psychological tools may differ from one culture to another because of appropriateness. Within the cultures and during cognitive education, children come across new psychological tools and they may develop some more general or specific cognitive functions over time.

During the development of children, they benefit from different kinds of mediations. These can be symbolic, as previously mentioned, or human-based (Kozulin, 2003: pp.15-38). Vygotsky (1978: pp. 90-91) defined human mediator in his theory, indicating that the activities firstly started in an environment where a child and an adult interact with each other and the child internalizes them as his psychological functions in time. Besides internalizing the functions, using mediation has several benefits on child development such as feeling a secure learning environment or encouragement and getting feedback (Schaffer, 1996: pp. 121-123). In her analysis of mediation, Rogoff (1995: pp. 139-164) emphasized three different aspects of mediation; namely, apprenticeship, guided participation, and appropriation. Apprenticeship is related to mediating socio-cultural patterns. Guided participation, as the name suggests, includes interpersonal features of the activities.

Last but not least, appropriation covers the changes that mediations cause during the activities. Moreover, teachers play an important role in mediating learners in the learning environments. They can mediate learners in terms of modeling, providing feedback, motivating, or cognitive structuring that is at a metacognitive level (Tharp & Gillimore, 1988: pp. 44-67).

In addition to teachers, parents can also mediate their children at home. Even in infants' development, caregivers and infants share an emotional communication developing through some stages (Lisina, 1986, as cited in Karpov, 2003: p. 142). According to Russian neo-Vygotskians, this emotional relationship between infants and caregivers is important for infants' development because the positive attitude that infants create affects everything and they see the caregivers as mediators in their relationship with the outside world (Kozulin, 2003: pp. 15-38). In this regard, Lehrer and Shumow (1997: pp. 41-83) investigated parental and teacher mediation while children were solving math problems. According to the results, they found that parents directly told children what to do whereas teachers tended to explain the problem and encouraged them to solve it themselves.

### 2.1.2. Regulation and Internalization

People begin to use other symbolic tools, such as language, over time, and they regulate themselves physically and psychologically by using symbolic tools (Poehner, 2008: p. 26). Regulation is defined as a form of mediation by Lantolf and Thorne (2006: pp. 197-221). During the cognitive development process, people follow some stages: object regulation, other regulation, and self-regulation (Bodrova & Leong, 2003: pp. 160-161). In the stage of object regulation, a child's psychological functioning depends on the environment rather than the individual (Poehner, 2008: pp. 27-28). Children use some objects to regulate themselves. To illustrate, they can look up the dictionary to learn new vocabulary. In the second stage, other regulation, they are regulated by other people around them. These people are generally more knowledgeable or skillful ones who guide and support them. As an example, students can ask their teachers or parents to learn something. Moreover, they can work collaboratively in a group (Poehner, 2008: p. 15). For Vygotsky, this

type of social interaction in which there is a cooperative or collaborative dialogue between a child and a more skillful interlocutor leads to cognitive development (McLeod, 2007: pp. 1-13). Swain (2000: pp. 97-114) also mentioned collaborative dialogue, emphasizing that it mediates language acquisition. In other words, both language use and language learning happen at the same time during collaborative dialogues.

There are also other researchers who highlighted the importance of interaction in language learning (e.g., Hatch, 1978: p. 401; Long, 1981: p. 259; Pica 1994: p. 493; Pica et al., 1987: p. 737). Some researchers among them such as Long (1981: p. 259) and Pica (1994: p. 493) emphasized "negotiation" during the interactions in learning environments. Pica (1994: p. 494) defined the negotiation as "modification and restructuring of interaction that occurs when learners and their interlocutors anticipate, perceive, or experience difficulties in message comprehensibility". In other words, interlocutors can repeat what they said, adjust the syntax, or modify the vocabulary depending on the understandability (Pica, 1994: pp. 493-527). Some researchers, on the other hand, emphasized the type and nature of feedback during these interactions (e.g., Aljaafreh & Lantolf, 1994: p. 465; Nassaji & Swain, 2000: p. 34).

After being guided by others and internalizing the information, children start to regulate their own performance (McLeod, 2007: pp. 1-13). At the level of self-regulation, individuals can mediate themselves through symbolic tools (Vygotsky, 1986: p. 171). They can control their actions through considering the alternatives and choose one of them intentionally (Poehner, 2008: p. 28). In other words, their mental processing relies on internal context more than external one; therefore, learning becomes more instrapsychological rather than interpsycological (Walsh, 2013: p. 8). In this way, Vygotsky (1994: p. 65) stated that through "internalization" or "ingrowing", individuals learn to use symbolic tools in the activities with others, resulting in a new form of cognition. Further, Vygotsky (1994: p. 66) explained the ingrowing process with a simple example that a child can remember the words given to them with the help of pictures in order after a few times and associate the words

with the pictures. Thus, this "complete ingrowing" process relies on inner stimuli instead of external stimuli (Vygotsky, 1994: p. 66).

On the way of development of self-regulation, preschool children speak to themselves to guide their behaviors or accomplish a task, which is called private or egocentric speech. Vygotsky believed that private speech is decreased and changed functionally (Bodrova & Leong, 2003: p. 160). In other words, after it is used externally in social life, it is internalized and becomes an inner speech rather than a public speech (Vygotsky, 1987, as cited in Bodrova & Leong, 2003: p. 160) and children use it as a psychological tool for self-regulation (Karpov, 2003: p. 145). In addition to private speech, Vygotsky also mentioned "make-believe play" which is another important context for self-regulation for children. Play is seen as a leading activity by Vygotsky (1977, as cited in Bodrova & Leong, 2003: pp. 161-162). During a play, children use some objects and actions. Since they internalize the cultural signs and symbols which are crucial in developing higher mental functions, Vygotsky (ibid.) thought that play helps children create their ZPD. A real play includes three features according to Vygotsky (1978: p. 90). Children firstly make up an unrealistic situation, determine some roles, act out, and obey some rules. In brief, make-believe play is important for developing the child's higher mental functions and promotes use of self-regulation.

### 2.1.3. Zone of Proximal Development

ZPD, which we encounter in many psychology and educational psychology books, was first introduced by Vygotsky. In his book "Mind in Society", he defined ZPD as "the distance between the actual developmental level as determined by independent problem-solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (Vygotsky, 1978: p. 86). It includes the process that a child will be able to accomplish a task with a more capable adult/peer that he will not be able to do alone. According to Vygotsky, learners internalize the social interactions and learning occurs in their ZPD (McCarthy & McMahon, 1992: p. 18). Chaiklin (2003: p. 41) also explained this process as "an interaction on a task between a more competent

and a less competent person where the less competent person becomes independently proficient at what was initially a jointly accomplished task". In other words, it is "what the child is able to do in collaboration today he will be able to do independently tomorrow" (Vygotsky, 1987, p. 211).

This concept is generally explained by comparison to Zone of Actual Development (ZAD) which is what a learner can do without help. In Vygotsky's (1978) expression, it is "the level of development of a child's mental functions that has been established as a result of certain already completed developmental cycles" (p.85) or "functions that have already matured, that is, the end products of development" (p. 86). In order to see the whole picture of the development of an individual, Vygotsky (1998: pp. 84-91) thought that only ZAD determined by individual performance is not sufficient because his potential ability surpasses his actual ability when someone more knowledgeable facilitates the learning. With a more knowledgeable other, the person has a chance to socialize and benefit from the mediation and assistance, which is important for cognitive development. Moreover, learner responsiveness to the external mediation such as interaction with an assessor or a peer as well as using models and charts gives information about still developing abilities of the learners (Poehner et al., 2014: p. 337). In learning environments including language learning, teachers and students construct meaning collectively. Now, the discussion turns to another important concept of SCT, i.e., scaffolding, which is very important factor in this process.

### 2.1.4. Scaffolding

Scaffolding is another mediational concept. At its simplest, it is based on the process where some linguistic support is given to the learners based on their needs during the learning process. Bruner (1978) firstly defined scaffolding as "... the steps taken to reduce the degrees of freedom in carrying out some task so that the child can concentrate on the difficult skill s/he is in the process of acquiring" (p. 19). Similarly, Donato (1994: p. 37) defined it as an environment in which a more knowledgeable person provides a supportive environment so that novice ones can extend their skills and knowledge. It is firstly used in a first language context as parents' interactional

support by Wood et al. (1976: p. 89). Learners sometimes need support before they can apply new skills and knowledge independently (Larkin, 2001: p. 30). Thus, scaffolding decreases gradually as learners internalize the new knowledge or complete the tasks. Moreover, this gradual decrease in teachers' support causes a gradual increase in students' responsibility (Rosenshine & Meister, 1992: p. 26).

Although Vygotsky did not use the concept of scaffolding, it is thought to be connected to his ZPD because both of them give particular importance to the assistance of a more knowledgeable person in a learning area. Wells (1999: p. 127) explained scaffolding as "a way of operationalizing Vygotsky's concept of working in the zone of proximal development". He also noted the important scaffolding features: "The essentially dialogic nature of the discourse in which knowledge is co-constructed; the significance of the kind of activity in which the knowing is embedded; and the important role played by the artifacts that mediate the knowing" (Wells, 1999: p. 127). Despite these ideas, some researchers think that scaffolding and ZPD are different in some aspects. For example, Shabani et al. (2010: p. 237-248) thought that scaffolding does not happen where the learner and the mediator collaborate and interact as in ZPD. Likewise, Daniels (2016: p. 59) stated that the scaffolding process includes one-way communication and the scaffolder provides mediation alone for the novice.

In the matter of the L2, we can see the effect of scaffolding in some corrective feedback studies which investigated error correction and language learning within learners' ZPD (e.g. Aljaafreh &Lantolf, 1994: p. 465; Nassaji & Swain, 2000: p. 34). For example, Aljaafreh and Lantolf (1994: p. 465) investigated the relationship of error correction with learning in the dialogic activities in which tutor and learners construct collaboratively, i.e., in their ZPD. Their results indicated that corrective feedback is effective as far as it is given in a interactional dialogue, negotiated by the teacher and learners, and provided in learners' ZPD. Therefore, Nassaji and Swain (2000: p. 35) thought that the study of Aljaafreh and Lantolf (1994: p. 465) is different from the conventional perspective because error correction was seen as a social activity where the teacher and the learners participated jointly.

In language classrooms, scaffolding is very important because it provides negotiation of meaning as well as linguistic support contributing to language development (Kayi-Aydar, 2013: p. 324). Some studies have supported the positive effect of scaffolding and scaffolded instruction on language learning (e.g., Huong, 2007: pp. 329; Kayi-Aydar, 2013: p. 324; Klingner & Vaughn, 2000: p. 69; Larkin, 2001: p. 30). Scaffolded instruction is defined as "the systematic sequencing of prompted content, materials, tasks, and teacher and peer support to optimize learning" by Dickson et al. (1993: p. 12). That is, teachers create an environment where support is systematically provided for the learners to accomplish the tasks. However, it not only includes teachers' support for the task accomplishment but also necessitates some systematic elements. In his study, Larkin (2001: p. 33) found that through scaffolded instruction, teachers make their students more independent learners and they become more responsible for their learning. Van Lier (2004: p. 151) mentioned about six characteristics of scaffolding in the field of language teaching: "continuity, contextual support, intersubjectivity, contingency, handover/takeover, and flow". Hogan and Pressley (1997, as cited in Larkin, 2001: p. 30) also indicated eight important components of scaffolded instruction: "pre-engagement; establishing a shared goal; actively diagnosing the understandings and needs of the learners; providing tailored assistance; maintaining pursuit of the goal; giving feedback; controlling for frustration and risk; and assisting internalization, independence, and generalization to other contexts". These are not for teachers to follow them compulsorily, but for them as a guideline for successful scaffolding.

## 2.2. LANGUAGE ASSESSMENT

Language assessment is as important as language teaching for understanding individuals' learning, knowledge, or ability in a language. Depending on the purposes, individuals are assessed in different ways. McNamara (2004: p. 56) stated that through assessments, we get information about learners' knowledge or ability. In addition to the tests in school settings, people may participate in language tests for their professional life, job employment, attending a course or a program, etc. McNamara (2000: p. 5) explained that these tests differ in terms of their purposes and methods. According to the method, tests can be broadly divided into traditional

paper-pencil and performance-based tests. Most of us are familiar with the traditional paper-pencil tests and they are generally used to assess separate language knowledge components (e.g., grammar, vocabulary, etc.) or receptive knowledge components (i.e., listening, reading). Moreover, the multiple-choice format is mostly used especially in standardized tests because it is easy and efficient to administer and score. On the other hand, in performance-based tests, practitioners mostly assess the speaking and writing skills of the test takers. They are also assessed in communication.

As for the test purpose, McNamara (2000: p. 6) indicated that tests mainly can be divided into two; achievement and proficiency tests. Proficiency tests measure individuals' ability in a language without taking into account the training or course they had. On the other side, achievement tests are related to the instruction and learning goals. They can be applied during (progressive achievement tests) or at the end of the course (final achievement tests) to evaluate the aims of the learning. On the other hand, Bachman (1990: pp. 60-61) explained another common distinction of assessments based on their purposes; summative and formative. Summative assessments are directly related to the results of the instruction and they occur at the end of the period to assess learners' success on specific content. Bachman (1990: pp. 60-61) indicated that summative assessments reflect the past learning of the individuals but the results are used to decide their futures. One of the most common summative tests is achievement tests because they assess learners' achievement on a specific content area. Bachman also explained that formative assessments are administered prior to the end of a course and their results provide important feedback for classroom teaching. Similarly, d'Anglejan et al. (1990) explained formative assessment which is used to gather

> information which will inform teachers and students about the degree of success of their respective efforts in the classroom. It allows teachers to diagnose students' strengths and weaknesses in relation to specific curricular objectives and thus guides them in organizing and structuring instructional material. (p.107)

Quizzes, planned tests as well as recordings of observations or informal teacher-student interactions, etc. are all examples of formative tests (Ellis, 2003: pp. 312-316). Based on the results of their study, Rea-Dickins and Gardner (2000: p. 229) also stated the features of formative assessment such as providing evidence of learning and instruction as well as helping teachers plan their teaching. All in all, Poehner (2007: p. 324) stated that many assessment types including traditional assessments (e.g., multiple-choice or open-response formats) and alternative assessments (e.g., portfolios or projects) have a summative or formative function in terms of the way their results are used.

Achievement tests include not only traditional tests such as the end of the course tests, but also tests based on the process of learning such as portfolio assessments. Like portfolios, the assessment types which are more innovative and reflect the progressive parts of the curriculum are known as alternative assessment. As the name suggests, it is an alternative to standardized tests and it is "designed to foster powerful, productive learning for students themselves" (Hargreaves et al., 2002: p. 70). Therefore, teaching and learning are integrated through alternative assessments. Naeini & Duwall (2012, p. 23) also explained the advantages of alternative assessment as both summative and formative, easy to interpret, and more informative. Although portfolios are the most popular and influential example of alternative assessments (Fox, 2017: pp. 135-148), there are also some others such as projects, observational checklists, diaries, self or peer- assessment, and poster presentation (Hargreaves et al., 2002: p. 74-75). In the next section, Dynamic Assessment, which also integrates teaching into the assessment from a social perspective, will be explained.

## 2.3. DYNAMIC ASSESSMENT

DA is an interactive alternative type of assessment combining assessment and instruction (Antón, 2012: p. 106). It is an approach developing rapidly and has been appeared in different forms and various contexts such as psychological, neuropsychological, educational, and language (Haywood & Lidz, 2007: p. 9). Considering the precedents of DA in the past, Poehner (2008: p. 23) stated that it can be associated with Socratic dialogues explained by Plato. Socrates creates an

environment including questioning and quick responses with his audience, and they create new perspectives together. Moreover, Poehner indicated that Socratic dialogues involve assessing and teaching at the same time, as in DA. In contrast to conventional tests, Socrates leads additional questions and suggestions to the audience to explore the topic in collaboration.

Vygotsky's SCT and especially his notion of ZPD provide a basis for DA. According to SCT, we are first mediated by physical and symbolic tools as well as through interactions with other people, and then we internalize these interactions, leading to new cognitive abilities (Poehner et al., 2017: p. 245). Through these social interactions, ZPD occurs and reflects development. As for assessments, Vygotsky was mostly interested in qualitative assessments and the dynamics of learners' development (Antón, 2012: p. 107). Moreover, he stated the need for diagnostic assessment focusing on future behavior rather than a symptomatic assessment based on current behavior (Minick, 1987: p. 116–40). Vygotsky also emphasized the actual and potential level of development. DA rejects the idea that a learner's success based on an independent performance (actual development) is predictive of his future performance (potential development); instead, DA focuses on the learner's responsiveness while being mediated in cooperation to predict his potential development empirically (Poehner et al., 2017: p. 245).

The term DA is not used by Vygotsky himself although he mentioned the implications of ZPD for assessing learners. However, the study of Luria (1961: p. 7), who is a close and influential colleague of Vygotsky, both precedes Vygotsky's works in English and pioneers all the works known as DA. Luria compared the assessment procedures using the term "dynamic" and he emphasized the importance of assistance during the assessments. Despite the importance of Luria's study, Budoff conducted the earliest DA research which was in English and received interest in education and psychology (Budoff, 1968: p. 295; Budoff & Friedman, 1964: p. 434). Like Budoff's studies, preliminary works in this field focused on children with abnormal behaviors, learning disabilities, or special needs. These studies include Keane's study (1987: p. 360) based on children with hearing problems, the study of Feuerstein, Rand, and Hoffman (1981: p. 465) about children with cultural deprival, and Samuel's study (2000: p. 521) on children with learning disabilities, etc. Besides

these research areas, researchers started to involve the concept of education and L2 in DA studies and the study of Lantolf and Poehner (2004: p. 49) brought L2 pedagogy into the field of DA.

In the field of L2, it is worth mentioning Aljaafreh and Lantolf‘s (1994: p. 465) ground-breaking research, which is a DA-focused. They investigated the impact of feedback and scaffolding on the development of students learning English as a second language (ESL). They developed a regulatory scale offering 13 types of feedback based on students’ ZPD when they make mistakes during the assessment. According to their regulatory scale (Aljaafreh & Lantolf, 1994: p. 468), assistance should be “graduated”, starting from the most implicit one to the most explicit one; it also should be “contingent”, meaning that the feedback should only be provided when it is needed.

### 2.3.1. The Definition and Characteristics of Dynamic Assessment

In the literature, DA is generally explained with an emphasis on its feature to incorporate teaching into the assessment procedure. For instance, Poehner (2008: p. 5) defines DA as “the unification of assessment and instruction grounded in Vygotsky‘s understanding of development”. The assessment refers to “understanding learners’ abilities” and the instruction means “supporting learner development” (Poehner, 2008, p. 2). Based on this idea, several researchers have attempted to define DA emphasizing this integration. To illustrate, Poehner and Lantolf (2005: p. 233) identified DA as “an approach to assessment and instruction derived from Vygotsky’s theory of the Zone of Proximal Development” and Stenberg and Grigorenko (2002: p. 23) also indicated that “Dynamic testing is testing plus an instructional intervention”. In that vein, Poehner and Lantolf (2005: pp. 237-238) highlighted that “As called for in Vygotsky’s ZPD, assessment and instruction are dialectically integrated as the means to move towards an always emergent (i.e. dynamic) future, rather than a fixed and stable steady state”. Furthermore, it should be noted that all educational activities should focus on learner development to fully integrate assessment and instruction (Poehner, 2008: p. 70).

DA is based on the idea that the abilities occur during the involvement in the activities where learners are supported by other people and with the help of available cultural tools (Poehner & Lantolf, 2013: p. 324). It includes the procedure that examiners provide active intervention, which is any kind of support given to the learners, and they assess learners' responses to this intervention (Haywood & Lidz, 2007: p. 26). In other words, assessment is not a separate activity and it is directly connected to the intervention (Lidz & Gindis, 2003: p. 100). Thus, the examiner plays an active role during the assessment and it occurs in an interactive manner between the examiner and the learners. Emphasizing these roles, Lidz (1987: p. 4) explained DA procedure as "an interaction between an examiner-as-intervener and a learner as- active participant, which seeks to estimate the degree of modifiability of the learner and the means by which positive changes in cognitive functioning can be induced and maintained". In this way, DA focuses on the intervention that helps learners move to the next competence level as well as the processes that explain observed behavior and often include meta-cognitive processes (Lidz, 1995: pp. 143-153).

Lidz (1991, as cited in Antón, 2009: p. 579) indicated that the activity by the learner and examiner, modifiability of the behavior based on mediation are important characteristics of DA. DA includes teaching through planned mediations and evaluates the effect of this teaching on learners' performance (Haywood & Tzuriel, 2002: pp. 40-63). During the assessment, examiners provide mediation so that learners overcome the problematic areas that they face; in this way, learners can perform beyond their current capabilities and the mediation leads to teaching aimed at supporting students' development abilities (Poehner, 2009: p. 472). In other words, the assessor or mediator intervenes when there is a breakdown in learners' performances through mediations such as leading questions, hints, or prompts. This process of mediation is not the same as corrective feedback and helps the assessors diagnose not only the learners' developed abilities determined by independent performance but also emerging abilities indicated by learners' responsiveness during this process (Poehner et al., 2017: p. 244). Taken together, DA procedure promotes students' development of certain abilities and provides insight into their potential for learning.

Dynamic Assessment has been also explained by its differences from statistical or non-dynamic assessments by some researchers. Statistic assessment generally refers to a procedure in which learners are presented with questions and rated without any intervention based on their responses to improve their performances (Tzuriel, 2001: p. 13). That is, standardized tests reflect learners' current level of development while DA involves mediation strategies and demonstrates the potential for learning. In his study, Luria (1961: pp. 1-16) compared the "statistical" and "dynamic" assessment approaches. To get information about a full picture of a person's capabilities, statistical assessment is based on the individual performance on a test while DA examines the performance with the assistance as well as the extent to which the assistance is used and the mediated performance is transferred to the other tasks.

In his paper, Luria (1961: pp. 1-16) investigated the children with learning disabilities in four groups in terms of their poor school performance: children with normal intelligence, but emotional issues, "weak children" with poor living conditions, and children having normal intelligence but "partial defects" such as hearing disorder (Luria, 1961: pp. 2-4). He also examined the way these children are placed into the school settings by Soviet educators and psychologists, and he stated that these children are not suitably distinguished. Moreover, when they are assembled in the institutions, they do not get enough and appropriate support for learning. That's why, Luria (1961: p. 5) criticized the traditional quantitative tests for measuring intelligence and suggested the idea that "the most important problem is that we have to pay more attention not only to the diagnosis but also to the prognosis of the developmental potential of these children". To illustrate the concept, he investigated three children, all of whom got the same IQ score on a traditional test considering only their independent performances. Luria stated that children are supposed to be assisted in the assessment based on ZPD and the "prognostic value" involves the analysis of not only how much children use the assistance but also the extent to which their performance improved with assistance. Last but not least, Luria (1961: p. 7) mentioned "the principle of transfer" that is testing the children later without assistance to get information about the improvement in their independent

performance. As a result, he highlighted the fact that not all children can benefit equally from the assistance and cannot maintain their improved performance. By emphasizing the importance of children's ZPD, Vygotsky explained this situation with a good example:

> Imagine that we have examined two children and have determined that the mental age of both is seven years. This means that both children solve tasks accessible to seven-years old. However, when we attempt to push these children further in carrying out the tests, there turns out to be an essential difference between them. With the help of leading questions, examples, and demonstrations, one of them easily solves test items taken from two years above the child's level of [actual] development. The other solves test items that are only a half-year above his or her level of [actual] development. (Vygotsky, 1956, as cited in Wertsch, 1985: p. 68)

Vygotsky also (1956, as cited in Wertsch, 1985: p. 68) indicated that "From the point of view of their independent activity they are equivalent, but from the point of view of their immediate potential development they are sharply different". Similarly, Luria (1961: p. 7) based on his study's results, concluded that "They [the three children in his example] may be quasi identical in a statistical approach, but they are not identical in a dynamic approach, in the zone of their potential development". That's why, assessing individuals' learning potential levels means "to create his or her ZPD through the interaction with the teacher/assessor" (Hessami & Ghaderi, 2014: p. 646).

Sternberg and Grigorenko (2002: pp. 28-29), like many researchers in the DA paradigm, also compared and explained dynamic and statistical assessments that include more traditional procedures. They stated that in statistical assessments, test items are presented to the examinees and they respond to them in a given time without any kind of feedback or intervention; instead, they only get the information about a report of the score(s) as feedback. On the other hand, Sternberg and Grigorenko (2002) explained DA as:

> a procedure whose outcome takes into account the results of an intervention. In this intervention, the examiner teaches the examinee how to perform better

on individual items or on the test as a whole. The final score may be a learning score representing the difference between pretest (before learning) and posttest (after learning) scores, or it may be the score on the posttest considered alone. (p. vii)

To sum up, DA is an umbrella term used for the heterogeneous approaches associated with a common feature of incorporating the instruction and feedback into the testing process (Lidz & Elliot, 2000: pp. 3-16). To assess learners' potential for learning, the teacher or the assessor provides mediation when learners encounter difficulty. During the assessment, the role of the teacher is basically observer, mediator, or collaborator. Furthermore, this mutual participation and interaction between the teachers and learners combine teaching and assessment (Poehner, 2009: pp. 471-472). Having demonstrated the general information and characteristics about DA, the next part will explain the approaches and formats to DA

### 2.3.2. Dynamic Assessment Approaches and Models

Besides the basic characteristics of DA, there are some differences among the approaches in it. In other words, there is no valid procedure in DA; instead, the approaches differ in terms of some aspects such as intervention and tasks used for the assessment (Lidz, 1995: p. 144). Lantolf and Poehner (2004: pp. 54-55) proposed two main approaches to DA; *interventionist and interactionist*. Both of them can be associated with different contexts where Vygotsky identified the ZPD: "Interventionist DA is rooted in Vygotsky's quantitative interpretation of the ZPD as a difference score" while interactionist DA "finds its origins in Vygotsky's second, qualitative, interpretation of the ZPD" (Poehner & Lantolf, 2005: p. 239). Moreover, unlike N-DA both of them focus on the development processes; however, they vary in terms of how researchers approach the mediation (Poehner, 2008: pp. 43-44).

In addition to the approaches, as for the positioning of mediation in the assessment process, Sternberg and Grigorenko (2002: p. 27) have proposed *sandwich* and *cake formats* of DA. Following standard experimental research design, the sandwich format involves a pre-test and post-test; the instruction is "sandwiched"

between these tests, which are given in a non-dynamic way (Poehner, 2008: p. 19). That is, learners take a static pre-test as the first step, and then the instruction, i.e., mediation phase, is provided to develop their skills in an individualized or group environment. Finally, they are administered the post-test in an alternative form of pre-test to analyze how much an individual made progress as a consequence of mediation.

On the other side, in the cake format, which is usually performed separately, a graded sequence of mediation is given from implicit to explicit when learners have difficulty in finding the correct answer for each item. Sternberg and Grigorenko (2002: p. 27) clarified that "the successive hints are presented like successive layers of icing on a cake", and the amount as well as the form of mediation depends on the learner while the content mostly does not change. In this way, the mediator can gain insight into how much support they need to find the correct answer during the assessment. All in all, the sandwich format provides a comparison of test performance before and after mediation while the cake format is considered effective so that learners can benefit from mediation whenever they need it (Sternberg & Grigorenko, 2002: p. 27). Therefore, both formats have different advantages in supporting learner performance through mediation.

#### 2.3.2.1. Interventionist Dynamic Assessment

In interventionist DA, a standardized process of administration and types of assistance is used to make comparison between or within the groups; in addition, the results can be used to compare with other measures and predict future performance (Poehner, 2008: pp. 44-45). In other words, the standardized mediations are predetermined for each test item before the test administration and they follow a scripted follow. These pre-scripted hints are offered to the learners based on a scale ranging from the most implicit one to the most explicit one (Lantolf, 2009: p. 360) because providing more explicit feedback when more implicit one is sufficient may prevent us from understanding the learner's level of development (Aljaafreh & Lantolf 1994: p. 467). Poehner (2008: p. 45) highlighted that the motivation behind using this move is that it increases the objectivity of the assessment and students'

performances can be evaluated according to some constructs of N-DA (e.g., validity, reliability, generalizability, etc.) through interventionist DA. Therefore, interventionist approaches are more in line with the assessments that are based on psychometric features rather than interactionist DA because the former provides measurable results of students' performance.

Budoff (1987: pp. 53–81; see also Budoff & Friedman, 1964: pp. 434–439) investigated the poor performance of learners in traditional intelligence tests and developed Learning Potential Measurement Approach that is based on traditional tests but includes a standardized mediational phase. He suggested that these tests may provide information about many learners' abilities while they might be inadequate to explain the abilities of those who are socioeconomically disadvantaged (Budoff, 1987, as cited in Poehner, 2008: p. 45). He stated that their poor performance may be affected by the culture of those learners or their schools, and insufficient learning opportunities rather than the impairments in their cognitive abilities (Poehner, 2008: p. 45). That's why, he believed that training about the test's familiarity and strategies on how to solve problems may help them develop their test performance (Sternberg and Grigorenko, 2002, p. 73) and his study follows and pioneers the sandwich format of DA including the design of pretest, treatment (where the mediation is provided for students as training about the strategies of problem-solving), and posttest. Although Budoff emphasized that this improvement in learners' performance shows their future learning potential (Budoff & Friedman, 1964: pp. 434-439), his main concern is to improve learners' test performance based on psychometric principles rather than developing them cognitively.

Guthke and his colleagues developed another interventionist DA approach which is Lernest Approach (Guthke, 1982: pp. 306-324.) or Leipzig Learning Test (LLT). Unlike Budoff's approach that includes mediation in the treatment phase, LLT provides mediation during the assessment itself and standardized hints follow five-step mediation (see Guthke & Beckman, 2000: p. 17) presented from implicit to explicit. Based upon the results of LLT, a score and a profile are presented for each learner; the former represents the number of mediation used and the time spent on the assessment whereas the latter refers to the error types and the form of mediation

that most leads students to the response (Poehner, 2008: p. 48). Learner profiles also shed light on the instruction phase because they reveal students' errors and responsiveness to mediation; therefore, assessment and teaching are integrated, which demonstrates the dynamic nature of LLT. Moreover, mediational hints are offered to learners in the next teaching environment within LLT and even the decrease in the number of mediation use in learners' performance unveils their development.

Similarly, as an alternative to LLT, Testing-the-Limits approach was developed by Wield and Carlson. During the assessment, they follow standardized mediations as in other interventionist DA approaches. However, they also encouraged learners to think aloud and verbalize their reasons during problem-solving so that the researchers can gain insight into the problematic parts and how learners approach the task for future instruction (Poehner, 2008: p. 49). By this means, more information can be obtained about the abilities of learners through this approach compared to other interventionist approaches such as Lernest and Learning Potential Measure (Poehner, 2008: p. 49). They also highlighted a consistent conclusion that students with poor performance in N-DA tests show more improvement than students with already high N-DA performance in terms of verbalizing (Kar et al., 1993: p. 26).

Graduated Prompt Approach (Brown & Ferrara, 1985: p. 273) that is developed by Brown and her colleagues follows standardized mediational phase in a sequence from implicit to explicit as in line with other interventionist DA approaches; however, they also included "transfer tasks" in the procedure of development based on Vygotskian perspective unlike aforementioned approaches (Poehner, 2008: p. 51). Transfer tasks shed light on whether learners can transfer skills developed with the help of mediation in their ZPD into new and more complex tasks. Besides informing about how far learners can transfer their skills and knowledge, transfer tasks also provide information on how much mediation they need for this. They gradually arrange the transfer tasks as near, far, and very far transfer tasks according to the degree of complexity and integration of familiar principles into the new tasks. Based upon the transfer task performance, they focused

on the speed of learning and the degree to which they extend the knowledge into new and more complex tasks.

#### 2.3.2.2. Interactionist Dynamic Assessment

Interactionist DA is based on the cooperative dialogues between the mediator and learner rather than the pre-determined levels of mediation. The mediation arises from these interactions during the assessment; hence it is highly dependent on the learner's ZPD (Poehner, 2008: p. 18). That is to say, the mediation in interactionist DA is more in accord with Vygotsky's insight of cooperation in ZPD and it follows a procedure starting with a more implicit form and becoming progressively explicit based on the learner's responsiveness to the mediation levels (Poehner et al., 2017: p. 246). Unlike interventionist DA that determines the mediation according to a "one-size-fits-all" approach, interactionist DA is based upon a "one-on-one format" (see Poehner& Lantolf, 2013: p. 325) and it is the most commonly used DA approach (Davin, 2013: p. 303).

Interactionist DA has its origins in Mediated Learning Experience (MLE) developed by Feuerstein. According to his Structural Cognitive Modifiability theory, cognitive abilities are not stable in our brain and they can be developed through communication and instruction (Feuerstein, Rand & Rynders, 1988: p. 48). Focusing on the modifiability of the learners, the theory explained direct learning, which is without mediation, and mediated learning. In direct learning where learners interact directly with the environment, Feuerstein associated the procedure of "trial-and-error" with the process of "stimulus-response" in Behaviorism (Poehner, 2008: p. 54). On the other hand, in mediated learning, a more competent person interacts with the learner, which Feuerstein called MLE, and provides assistance through mediation. In this regard, Feuerstein demonstrated DA including the procedure that the mediator provides prompts and assesses learners' responsiveness to these prompts at the same time to identify their potential abilities. In his analysis of MLE (Feuerstein, Rand & Rynders, 1988: p. 20), Feuerstein explained 11 features of MLE emphasizing the first three; "intentionality and reciprocity, transcendence, mediation of meaning". From the Vygotskian perspective, MLE enables these interactions to be

internalized and it increases learners' capacity to make use of direct learning exposure (Feuerstein, Rand & Rynders, 1988: p. 58).

Needless to say, both approaches have different strengths and limitations. For example, as mentioned above interventionist DA might be more suitable for larger groups while interactionist DA might be more helpful in classroom settings (Antón, 2012: p. 108). Furthermore, interactionist DA does not have a high capacity to include different components of a structure in a single study (Kamrood et al., 2019: p.5). In this regard, Thouësny (2010: p. 3517) listed some differences between the two approaches based on the related literature. Apart from the features mentioned above, Interventionist DA is suitable for both individual and group settings while interactionist DA is practiced individually, thus it is more time consuming when compared to the interventionist approach. About the manner of mediation, in the interventionist DA, mediational prompts can be provided both in written and spoken whereas only spoken mediation is offered to the learners in interactionist DA because of its dialogic feature. From the point of validity, reliability, and generalizability; interventionist DA is more in line with the N-DA (Poehner, 2008: p. 45). Due to the above-mentioned limitations of interactionist DA, researchers such as Lantolf and Poehner (2008: p. 177) suggested interventionist C-DA which will be further clarified in the following sections.

### 2.3.3. Dynamic Assessment and Second Language Research

DA is a field that is encouraging, but still developing within educational psychology (Grigorenko, 2008: p. 127). Furthermore, it is also an area of increasing interest in L2 learning and teaching pedagogy (e.g., Antón, 2009: p. 576; Lantolf & Poehner, 2004: p. 49; Poehner & Lantolf, 2005: p. 233; Poehner, 2005: pp. 314-327; 2009: p. 471; Ableeva, 2008: p. 57; Summers, 2008: pp. 360-362). However, it is a relatively new area in applied linguistics (Poehner, 2008: p. 91), thus its effect and potential in language learning still need to be explored. By providing valuable information on the theoretical principles of DA to applied linguistics and encouraging the use of DA as a pedagogical method, Lantolf and Poehner made a significant contribution to the field (Ableeva, 2010: p. 130). Further, some

researchers use the term “dynamic testing” (e.g., Sternberg & Grigorenko, 2002: p. 23; Wiedl, 2003: p. 93) or “learning potential assessment” (e.g., Budoff, 1987: p. 53) in their studies.

From the Vygotskian perspective, DA is investigated as a procedure that not only reveals the developed abilities of learners but also informs about the developing functions and puts their future developments in the foreground in L2 research (Lantolf & Poehner 2004: p. 54). Integrating assessment and instruction, DA provides graduated mediation for learners when they encounter difficulties during the assessment so that they can immediately overcome the problematic parts of their performance as well as internalize and transfer the knowledge for future tasks. Within L2 DA research, there are studies following both interactionist (e.g., Antón, 2009: p. 576) and interventionist approach (e.g., Kozulin & Garb, 2002: p. 112). Moreover, some studies have been conducted to investigate the impact of DA on various areas of language learning including vocabulary (e.g., Hessamy & Ghaderi, 2014: p. 645; Veen et al., 2016: p. 329; Zadeh, 2018: p. 1), grammar (e.g., Malmeer & Zogh, 2014: p. 1707; Sharafi & Sardareh, 2016: p. 102), listening (e.g., Ableeva, 2008: p. 57; 2010: pp. 356-361; Hidri, 2014: p. 1; Wang, 2015: p. 1269), speaking (e.g., Siwathaworn & Wudhayagorn, 2018: p. 142; Yılmaz-Yakışık, 2012: pp. 129-136), and reading (e.g., Ajideh & Nourdad, 2012: p. 141; Guterman, 2010: p. 283; Naeini & Duvall, 2012: p. 22). Before moving on to the primary concern of this research, it is necessary to provide some important and exemplary studies in L2 DA pedagogy that focus on different language aspects.

Early DA studies in L2 context include investigations of L2 students at risk (e.g., Kozulin & Garb, 2002: p. 112), such as those with dyslexia (e.g., Schneider & Ganschow, 2000: p. 72). In their major study, Kozulin and Garb (2002: pp. 112-127) examined the text comprehension of 23 at-risk students in terms of academic education in Israel. Following an experimental research design and interventionist DA approach, students were administered a non-dynamic pre-test, mediational phase aimed at teaching EFL reading comprehension strategies, and another non-dynamic post-test to investigate students’ ability to master and use the strategies. The

mediation phase sought to improve post-test performance of students and to promote development. They also attempted to analyze students' abilities by measuring a single score, Learning Potential Score (LPS), instead of detailing the students' performances during the phases qualitatively (Poehner, 2008: p. 94). Kozulin and Garb (2002: p. 121) stated that "the learning potential score (LPS) has to reflect the gain made by the student from pre-test to post-test and an absolute achievement score at the post-test" and they developed a formula to calculate LPS to differentiate students with high and low learning potential, as seen in Figure 1.

**Figure 1**

*Learning potential score (LPS) formula (Kozulin &Garb, 2002: p. 121)*

$$\text{LPS} = \frac{(\text{S post} - \text{S pre})}{\text{Max S}} + \frac{\text{S post}}{\text{Max S}} = (2\ \text{S post} - \text{S pre})/\text{Max S}$$

*Note. LPS: learning potential score; S: score.*

Based upon the results, they classified the students into low, mid, and high LPS groups. The findings revealed that students with the same pre-test scores differ in terms of their post-test scores and hence LPS. Correspondingly, they also demonstrated the effect of DA in revealing the information about students' learning potential that could not be obtained with static tests as well as the fact that those students have different learning needs. Additionally, they highlighted the importance of some aspects that we should be careful while interpreting the results of DA procedure generally and LPS scores. Firstly, they noted the limitations of generalizability of DA results and reliability of LPS because the extent to which the mediation was qualified depends on the intervention provided by the examiner, which means that a different mediator may obtain different results on learning abilities from the same group of students. Secondly, they remarked that students' content knowledge (e.g., vocabulary, grammar) may also affect their performance in language tasks. As a result, they suggested individualized plans and instructions for future learning for students with different learning needs in light of their learning potential.

Following the interactionist DA approach, Antón, (2009: p. 576) integrated the DA procedure into a university-level advanced Spanish language program to examine learners' abilities through intervention and report their development. Learners are administered a diagnostic test including five parts: Traditional grammar-vocabulary, listening comprehension, and reading comprehension tests; dynamic writing and speaking tests. The findings demonstrated that DA displays both current and developing abilities of learners simultaneously although the interactionist DA takes a long time and a lot of effort, and may probably difficult to use in large groups. As a result, she suggested that these qualitative analyses of learners shed light on the programs in terms of preparing future learning plans based on their individual needs (Antón, 2009: p. 591).

Poehner (2005: pp. 314-327), in his doctoral dissertation, examined the oral abilities of six undergraduate learners in L2 French to analyze and improve their abilities provided by DA. He also sought to find out the benefit of DA in guiding individualized learning instruction according to their ZPD, and to what extent the development is not limited to specific assessment contexts. Firstly, learners were administered pre-tests in both static and dynamic procedures and based on these results the participants were included in the six-week enrichment program involving individualized sessions with the researcher. The sessions, in which learners were asked to narrate short video clips verbally in French, aimed to enable students to use "passé compose" and "imparfait" appropriately and specify the problems observed during the assessments. Finally, learners took the post-test and transfer test after the enrichment program. The overall analysis of the results suggests that DA is an efficient way of recognizing the abilities of learners and assisting them to deal with language difficulties they face.

As mentioned above, the applicability and advantages of DA over traditional static assessments in different skills in L2 contexts have been explored by the researchers, especially since the second millennium. In addition, L2 DA research is not limited to English language; instead, there are also other DA studies conducted in different contexts (e.g., French and Spanish). For example, to analyze the effect of DA on listening ability and compare it with traditional tests; Ableeva (2010: pp. 356-

361), for his dissertation, examined the improvement of university students who learn French as an L2. As a guide for her dissertation, she conducted a pilot study (i.e., Ableeva, 2008: p. 57), which revealed that the DA procedure is more explaining the differences in determining the source of the problems students encountered in listening comprehension rather than non-dynamic tests (Poehner, 2008: p.97). In the research (Ableeva, 2010: pp. 356-361) following an interactionist DA approach to cooperate with the students on an individual basis, students took non-dynamic and dynamic tests as well as enrichment program including authentic listening materials such as interviews and radio commercials. By offering mediations based on the graduated prompt approach (see Brown & Ferrara, 1985: p. 273), the students were also administered transfer tasks with increasing difficulty, which provides them more learning opportunities to sustain their abilities gained through DA sessions to other tasks as well, that is, promote development. As a result, it has been concluded that it is possible to assess students' actual level and to diagnose their potential level of development in listening comprehension through DA, which offers an insight into further instruction and intervention for students to handle problems (Ableeva, 2010: pp. 365-366).

Regarding vocabulary learning in L2, Ebadi et al. (2018: p. 1) recently investigated students' lexical inferencing in two groups through different assessment tools, i.e., static test and C-DA. The students in the C-DA group were provided graduated hints when they had difficulty in inferencing the meaning of unknown words, unlike the static assessment group. According to the results of the pre and post-tests, it was found that students in the C-DA group were better at gaining vocabulary than those in the SA group, which supports the idea DA promotes improvement in vocabulary acquisition through lexical inference. More recently, Ebadi and Rahimi (2019: p. 1) examined the short and long-term effect of online DA on EFL university students' academic writing skills. Students were supported with online generated mediation during synchronous writing sessions via Google Docs. The findings indicated that students had some problems in transferring their abilities into more challenging writing tasks; however, online DA sessions have benefited EFL teachers in greatly improving the academic writing skills of EFL students, and DA enables learners to reach a level of potential functionality in academic writing

skills (Ebadi et al., 2018: p. 24). Following Group Dynamic Assessment (G-DA), which was developed as a solution to the criticism of the interactionist DA approach (see Poehner, 2009: p. 471), Mehri and Amerian (2015: p. 11) investigated both the practicality and effect of G-DA on the development of students with different proficiency levels in past tense through transcendence writing tasks. According to their findings, DA sessions provided significant improvement in students' control of past tense in writing.

Some researchers in Turkey, as in other countries, have also attempted to explore the effect of DA on various L2 aspects of students at different proficiency levels and in different contexts. However, the effect and applicability of DA in L2 environments have not been widely practiced in the Turkish context. For instance, there are some studies analyzing the impact of DA on speaking ability (e.g., Çetin-Köroğlu, 2019: p. 23; Yılmaz-Yakışık, 2012: pp. 129-136; Yılmaz-Yakışık & Çakır, 2017: p. 22), speaking ability and metacognitive awareness (e.g., Kır, 2020: pp. 103-105), grammar preference in writing (e.g., Şentürk, 2019: pp. 72-74), metasyntactic awareness (e.g., Çalış, 2018: pp. 89-96), use of metadiscourse markers in writing (e.g., Ulu, 2020: pp. 98-99), and agency of learners (e.g., Güleryüz-Adamhasan, 2019: pp. 124-129). Although their context and methodology are different from each other, they have similar results revealing the positive effect of DA on learners' development in the area of research, and some of them also demonstrated learners' positive attitudes towards DA. Nevertheless, despite the positive effect and advantages of DA over traditional assessments, further investigation is required to confirm the findings. Moreover, for L2 learners, reading comprehension is a crucial ability for general language learning, but DA research on L2 reading comprehension is also quite limited.

### 2.3.4. Computerized Dynamic Assessment (C-DA)

In recent years, there has been a growing interest in applying the basis of SCT in computerized L2 research (e.g., Lee, 2008: p. 53; Li & Zhu, 2013: p. 61; Ma, 2017: p. 183) and DA researchers also have started to integrate computerized mediation into assessment procedure including several language skills. Among DA

researchers, Lantolf and Poehner (2008: p. 177) suggested interventionist C-DA as a solution to the aforementioned limitations of interactionist DA. Like other approaches and models of DA, C-DA has advantages and disadvantages compared to ordinary DA. For instance, Poehner (2008: p. 177) listed the important advantages of C-DA over DA as follows "… it can be simultaneously administered to large numbers of learners; individuals may be reassessed as frequently as needed; and reports of learners' performances are automatically generated". There are, however, also some criticisms of C-DA. To illustrate, C-DA is limited in proving individualistic mediation, namely based on learners' ZPD like other interventionist DA approaches. Since the prompts are pre-scripted and the same for all learners, they may not be sufficient for learners' needs although they are structured elaborately (Poehner, 2007: p. 326).

Even though most of the C-DA studies (e.g., Poehner & Lantolf, 2013: p. 323; Poehner et al., 2014: p. 337) have revealed the potential of CDA to learn about both developed and potential abilities of learners, L2 C-DA research has been an area still not fully explored, but a growing area of research. Guthke and Beckmann (2000: p. 17) firstly implemented computerized mediation into "Lerntest" (Leibzig Lernest test) which was structured to assess the cognitive abilities of learners through prompts (e.g., language aptitude). This computerized version of LLT includes a procedure as follows; if students respond incorrectly to the items, training exercises involving relevant concepts and examples are presented to the learners and they retry the items accordingly. More precisely, the computerized version of LLT defines where difficulties are faced by learners and offers support through training tasks for them to master the elements including complex items and to progress with more challenging questions during the assessment (Poehner, 2008: p. 178). As for interpreting the results, the items that they need a tutorial to respond correctly represent their ZAD while the ones that they answer correctly with the help of tutorials reveal their ZPD (Poehner & Lantolf, 2013: p. 326).

Another preliminary work on C-DA was undertaken by Jacobs (1998: p. 113; 2001: p. 217) through the program known as KIDTALK (Kidtalk Interactive Dynamic Test of Aptitude for Language Knowledge). As the name of the test

suggests, it was designed to assess the language aptitude of children in pre-school and school-age through computerized activities (Poehner, 2008: pp. 177-178). Children are firstly presented the training session including videos with puppets designed to teach the vocabulary and morphology of an invented language which is "kidtalk" and they are asked to answer the following questions about understanding the language (Poehner, 2008: p. 178). Similar to the procedure of computerized LLT, when learners give an incorrect answer, the relevant video part is displayed, and then they reattempt the item. In addition, training videos are presented until learners respond incorrectly on their third attempt; otherwise, the program leads them to the next item if they still do not answer correctly after the third attempt. At the end of the test, a report representing the scores based on the correct answers, regardless of the number of attempts, and another report based on the number of attempts are provided for each child (Jacobs, 2001: p. 224).

Although C-DA studies mostly adopt the principles of interventionist DA because the computer provides mediation, some researchers also integrate the interactionist approach into the C-DA procedure with the help of a human mediator. As an important example, Tzuriel and Shamir (2002: p. 23), in their research in cognitive psychology, designed a C-DA to evaluate kindergarten children's serial thinking skills, which are related to their performance in mathematics in their opinion. In their C-DA approach, children are asked to distinguish among a series of shapes based on their features (i.e., size, number, or darkness) and according to which criteria they will classify varies for each task (Tzuriel & Shamir, 2002: pp. 24-26). Following an experimental design to investigate the effect of C-DA, they assigned the learners into a group provided computer-assisted mediation and another group supported by only an examiner. If the children in the C-DA group give incorrect answers, a series of mediation from implicit to explicit is offered by the computer program; meanwhile, the examiner also participates in the assessment procedure and provides assistance when learners need additional support. As a result, the test procedure, in which an examiner is also included in the testing phase in addition to the computer-assisted mediation, found to be more effective in promoting cognitive changes than the procedure in which only an examiner is involved (Tzuriel & Shamir, 2002: p. 21).

#### 2.3.4.1. C-DA Research on L2 Reading Comprehension

C-DA is viewed as an alternative solution to the concerns of especially interactionist DA approach in terms of practicality and learner/sample size. However, there are relatively few studies on CDA in the field of second or foreign language learning. Since the primary concern of this research is C-DA and students' reading skills, the following review will mainly focus on C-DA studies in relation to reading comprehension in the field of L2 learning or assessment research. Before starting, it should be noted that the study of Aljaafreh and Lantolf (1994: p. 465), which suggests a regulatory scale providing graduated prompts and hints to learners based on their ZPD. Following their study, most of the interventionist C-DA studies have benefited from their regulatory scale to develop appropriate mediation according to learners' ZPD.

Poehner and his colleagues conducted early and important studies on CDA implementation in L2 context, which paved the way for other L2 C-DA research (e.g., Poehner et al., 2014: p. 337; Poehner & Lantolf, 2013: p. 323). Poehner and Lantolf (2013), in their impressive study, examined L2 reading and listening development of students learning Chinese and French as an L2 through interventionist C-DA. The test includes DA items related to reading and listening comprehension as well as transfer items, which are more challenging, in the multiple-choice format. The purpose of including transfer items is to find out whether students can benefit from their knowledge in other questions, varying in difficulty in different contexts. Regarding the preparation of appropriate mediation, they organized the prompts based on a pilot study involving one-to-one interactions between the mediator and several students. Students' performances in the test were generated with a different scoring system as three different scores: actual score representing their performance without mediation based on learners' first attempt (either four or zero points), the mediated score calculated according to their responsiveness to mediation (mediated score depends on the number of attempts and it decreases if the number of try increases), and LPS calculated with the formula of Kozulin and Garb (2002: p. 121), focusing on the gain between actual and mediated scores and thus providing information about the potential of learners (Poehner &

Lantolf, 2013, pp. 331-332, see Figure 2). In addition, they mentioned and calculated the "gain score" reflecting "the change between the actual and mediated components of the tests" and their findings indicated the significant differences between actual and mediated scores from Chinese listening and reading; French listening tests, as evidenced by the gain scores (Poehner & Lantolf, 2013: p. 334). Moreover, the analyses of the transfer items demonstrated that the learners with similar actual scores differ in their mediated performances, reflecting their LPS differing from each other despite the same unmediated performance. Based on the actual scores obtained from the transfer items, Poehner and Lantolf (2013: p. 336) indicated that "LPS has promise as a predictor of learning".

**Figure 2**

*The Formula of learning potential score (LPS) used by Poehner and Lantolf (2013)*

$$LPS = \frac{(\text{S post-mediation} - \text{S actual})}{\text{Max S}} + \frac{\text{S post-mediation}}{\text{Max S}}$$

Abbreviated as follows:

$$LPS = \frac{2\,(\text{S post-mediation}) - \text{S actual}}{\text{Max S}}$$

*Note. LPS*: learning potential score; S: score.

Unlike Poehner and Lantolf (2013: p. 323), Poehner et al. (2014: p. 337) focused on how the results from a C-DA including L2 Chinese reading and listening comprehension items can be used as a tool to diagnose students' development in the context of L2. The standardized mediations they organized for C-DA came from the pilot studies in a non-dynamic and an interactionist DA format. Similar to Poehner and Lantolf (2013: p. 323), the scores of actual, mediated, transfer and LPS were calculated for each learner and the results showed the significant difference and correlation between learners' actual and mediated scores. Moreover, they found a significant negative correlation between the gain and actual scores, which reflects that learners with lower independent scores benefit from the mediation more than those with better independent scores (Poehner et al., 2014: pp. 337- 357). As

expected from the gain scores, learners with the same actual scores have different both mediated and learning potential scores. While interpreting the data, they also investigated the specific language areas that learners face difficulty. Based on the overall results, they suggested that the scores earned from the test not only diagnose learners' progress but also provide insight into future teaching based on their needs.

Differently from most of the C-DA studies implementing one-shot C-DA, Teo (2012: p. 10), in his action research, included C-DA in an eight-week enrichment program between pre and post-test following sandwich format of interventionist approach to help Taiwanese university EFL learners improve their inferential reading skills. During the mediational phase, learners were provided prompts from implicit to explicit and the findings of pre and post-tests displayed a significant improvement in learners' inferential skills. The learners also stated the usefulness of CDA in promoting their metacognitive reading strategies in written reflections in their portfolios.

Yang and Qian (2017: pp. 1-15) also explored the effect of C-DA, learners' reflections about C-DA, and the difficulties that learners had during the reading comprehension test. Through C-DA including multiple-choice reading comprehension questions supported with graduated mediational prompts, they used the scoring procedure of Poehner and Lantolf (2013: pp. 331-332); that is, the program offers learners' actual, mediated, gain scores as well as LPS. According to the results, learners reported that they mostly had difficulties because of the unknown vocabulary, finding the relevant part of the passage, and their poor inference skills respectively (Yang & Qian, 2017: p. 10). Secondly, a significant difference was found between learners' actual and mediated scores as well as the positive and strong correlation between these two scores. Furthermore, they discovered a negative correlation between their actual and gain scores, indicating that learners with a poor independent performance benefit more from mediation than those with higher actual scores. Their findings also showed that the learners with the same actual scores differ in their mediated scores and therefore LPS. Finally, the questionnaire results revealed that most of the learners, especially low achievers, believed that C-DA was useful and helpful for the improvement of their reading comprehension skills.

More recently and following a quasi-experimental design, Yang and Qian (2019: pp. 1-15) examined the reading comprehension abilities of Chinese EFL learners through integrating C-DA into a four-week enrichment program. While the learners in the control group took the traditional teaching method and multiple-choice test, the C-DA was employed in the experimental group as an enrichment and assessment tool. The results obtained from pre, post, and transfer tests demonstrated that learners in the C-DA group took advantage of the intervention while the control group was unable to make progress in the post-test; in addition, although the performances of learners in both groups decreased in the transfer test, the ones in the C-DA group still had a much better transfer scores than those in the N-DA group (Yang & Qian, 2019: p. 16).

When the recent studies on C-DA are examined, it can be observed that most of the studies were carried out in the Iranian context both related to reading (e.g., Barabadi, 2010: p. 1; Ebadi & Saeedian, 2015: p. 1; 2016: p. 27; 2019: p. 51; Estaji & Saeedian, 2020: p. 347; Pishghadam & Barabadi, 2012: p. 73; Pishghadam et al., 2011: p. 1353; Shabani, 2012: p. 15) and other skills such as listening (e.g., Kamrood et al., 2019: p. 1), writing (e.g., Davoud & Ataie-Tabar, 2015: p. 176), and grammar (e.g., Ahmadi & Barabadi, 2014: p. 161). For instance, Barabadi (2010: p. 1) implemented a computerized dynamic test (CDRT) to examine Iranian EFL learners' reading comprehension. The results of the study illustrated that graduated hints related to reading strategies contributed greatly to the learners' reading comprehension. Similarly, Pishghadam et al. (2011: p. 1353) investigated the impact of C-DA on reading comprehension of Iranian EFL learners who were offered scripted mediations from implicit to explicit during the assessment. The results presenting both non-dynamic and dynamic scores revealed that the C-DA procedure made an important contribution to the learners' reading comprehension scores as evidenced by the significant difference between the two scores mentioned. Further, they noted that based on their non-dynamic test scores, low achievers benefited more from mediation than high achievers. Pishghadam and Barabadi (2012: pp. 73-95) also found similar results in their study aimed at creating and validating a C-DA and examining its effectiveness in improving reading skills, indicating that there was a significant difference between non-dynamic and dynamic test scores of learners.

They indicated that learners' non-dynamic and dynamic scores were highly correlated while the correlation between their non-dynamic and gain scores dropped drastically. Moreover, LPS results clearly illustrated the extent to which learners benefited from mediation and made progress during the test and that the learners with the same non-dynamic scores may have different LPS.

Ebadi and Saeedian (2015: p. 1; 2016: p. 27; 2019: p. 51) also found similar results in line with the above-mentioned studies regarding the difference between the N-DA and C-DA scores of learners and the diagnostic feature of C-DA in revealing learners' abilities. Ebadi and Saeedian (2015: pp. 1-26) focused on the development of reading skills of at-risk Iranian EFL learners with the help of implementing C-DA created by Pishghadam and Barabadi (2012: pp. 73-95). Following the research design including a non-dynamic pre-test, enrichment program involving DA sessions, and a post-test (CDRT), they found that learners' performances increased significantly in the post-test after the enrichment program. Likewise, Ebadi and Saeedian (2016: p. 27) also integrate transcendence (transfer) assessment in addition to pre-test, post-test, and enrichment program to analyze whether learners can expand their growth into more demanding and novel contexts. They concluded that obtained scores indicated both learners' performance maintenance and development in reading ability. Ebadi and Saeedian (2019: pp. 51-78.) more specifically concentrated on the revealing effect of C-DA in learning potential of reading ability and they concluded that learners who gained the same pre-test scores may have different DA post-test scores and therefore different LPS emphasizing that non-dynamic tests are inadequate to inform about learners' potential for learning.

The C-DA studies so far have provided useful information about actual and mediated scores of learners that reveal their ZPD. This, in turn, has provided stakeholders with information about future teaching in line with learners' needs. However, more investigations are required to examine the effectiveness of C-DA, especially in Turkey. Since primary aim of this research is to find out whether students' WM scores predict their N-DA and C-DA scores, the following review will explain the studies on the importance of WM in L2 reading comprehension.

## 2.4. DEFINITION AND MEASURES OF WORKING MEMORY

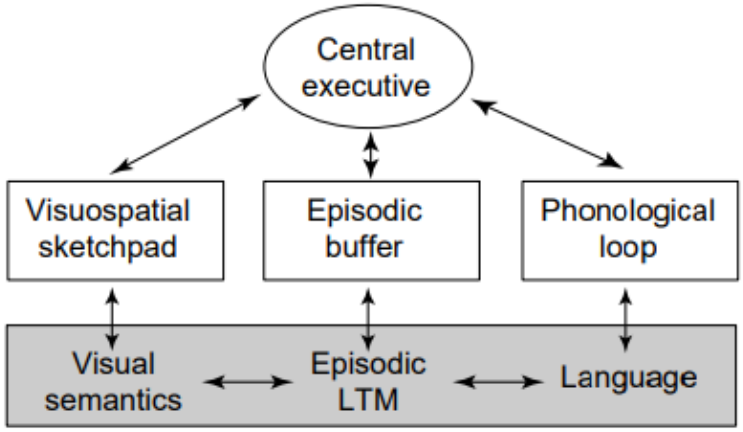
WM, a system with a restricted capacity, takes an active role in storing temporary information during the cognitively challenging tasks (Baddeley & Hitch, 1974: pp. 47-48). In other words, it is dynamic in nature because it is involved in the processes that can be distracting because the information is controlled, organized and, actively stored when other information is also available (Conway, Jarrold et al., 2007, as cited in Chang et al., 2019: p. 1). The role of WM, therefore, is critical in information processing (Baddeley & Hitch, 1974: p. 85). WM is often confused with short-term memory (STM) and some use them interchangeably. However, it is possible to differentiate them in many ways even though both of them have limited capacity to store. To begin with, STM takes part in the situations where individuals are not supposed to make an inference or modify requirements of the process; instead, they are expected to remember the items in the correct order presented when measuring their STM (Swanson, 2006: p. 126). Moreover, it is seen as a basic storage system where its capacity depends on the practice through "rehearsal and chunking" (Conway, Cowan, et al. 2002: p. 164). On the other hand, it is believed that WM has both processing and storage functions as well as a specific function associated with transferring the information to the long-term memory (Baddeley & Hitch: p. 86). For example, participants need to remember some necessary task items and also ignore the others during WM tasks (Swanson, 2006: p. 126). All in all, WM represents a more active part of the system of human processing while STM is generally viewed as a buffer including passive storage (Newell, 1973, as cited in Daneman & Carpenter, 1980: p. 450).

Among the models developed for explaining WM (e.g., Cowan, 1988: p. 163; Engle et al., 1999: p. 309), the cognitive model of Baddeley and Hitch (1974: p. 47) is the most common and influential model. In their model, which is presented in Figure 3, there are two slave components which are the *phonological loop* associating with language and control of behavior, and the *visuo-spatial sketchpad* integrating into a single representation of visual, spatial, and probably kinesthetic information; and the *central executive system* which two mentioned sub-systems depend on and controls the attention processes (Baddeley, 2003: pp. 199-200). The

central executive component also plays a vital role in controlling the reading processes where individuals make inferences, interpret, or modify the information in WM (Miyake et al., 2000: p.51). Further, Baddeley (2000: p. 421) proposed the fourth sub-system, the episodic buffer, which encodes the information gathered from the phonological loop and visuo-spatial sketchpad to the long-term memory.

**Figure 3**

*Baddeley's Revised Working Memory Model (2000: p. 421)*



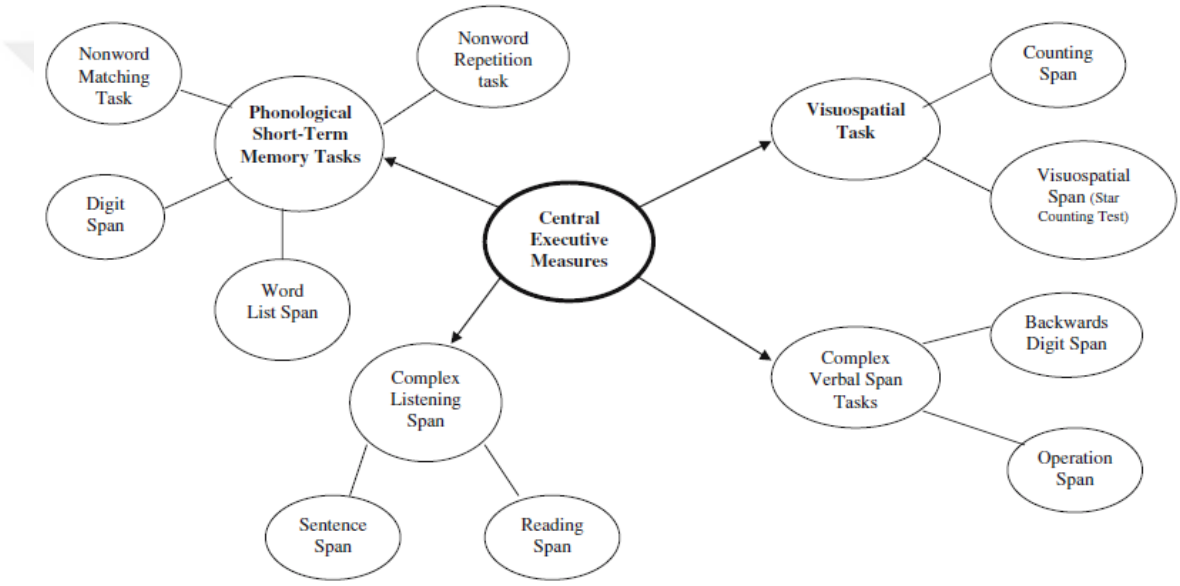

*Note.* LTM: long-term memory

Before starting a review of the relevant literature, it is worth mentioning the various measures used for especially verbal WM. To measure WM, as presented in Figure 4, researchers have used different kinds of tasks including the functions of processing plus storage (e.g., reading and listening span tasks); and storage alone tasks, namely traditional simple span tasks (e.g., word span and digit span tasks). Simple span tasks require the functions related to simple storage. For example, in the digit span task (Yuill et al., 1989: pp. 351-361), participants are expected to read aloud the digit groups and remember the last digit at the end of each group. On the other hand, *complex span tasks* are widely used especially to measure learners' verbal WM. They also differ in themselves in terms of the task used and the focused skill/ aspect such as listening span and reading span task. They involve both processing and storage functions because the integrated effect of processing and storage, rather than the effect of either, is better at explaining the variations in comprehension (Savage et al., 2006: p. 191). For this reason, tasks that necessitate

both storage and processing functions have been suggested to measure WM as a reading comprehension predictor, rather than traditional STM tasks such as word span and digit span (Daneman & Carpenter, 1980: p.450; Daneman & Merikle, 1996: p. 422; Linck et al., 2014: p. 874). Furthermore, the *reading span task (RST)* developed by Daneman and Carpenter (1980: p. 450) is the most preferred task especially when it comes to reading comprehension.

**Figure 4**

*Different Working Memory and Short-term Memory Tasks (adapted from Savage et al., 2006: p. 189)*




In contrast to RST, Turner and Engle (1989: p. 127) measured WM using mathematical operations rather than sentences through the *operation span task.* Shah and Miyake (1996: p. 4), as another complex span task, developed the *rotation span task* where learners are expected to decide whether rotated letters are normal or mirror images. As another example, in the *counting span task* developed by Engle et al. (1999: p. 309-331), participants are presented dark blue circles and squares as well as other circles in different colors; they are supposed to count the dark blue circles among the figures while they also need to retrieve the number of circles in the order presented in the series. Kane et al. (2004: p. 189) developed the *symmetry span task* and *the navigation span task*. As a part of the visuo-spatial complex span tasks, the former focuses on deciding symmetry and remembering the location of colored

cells in the matrices while the latter involves mentally moving an asterisk on a letter and retrieving the paths of a moving ball.

## 2.5. THE ROLE OF WORKING MEMORY IN LEARNING

The predictive power and relationship of WM in different fields have been researched for years and it is considered a good predictor of overall academic achievement (Holmes et al., 2009: p. 9). WM is also thought to contribute to children's learning, school success (Alloway & Alloway, 2010: p. 20), and language development (Weismer et al., 2000: p. 865). However, its role in language comprehension has been particularly emphasized because individuals process a set of symbols during comprehension (Just & Carpenter, 1992: p. 122). For instance, Daneman and Carpenter (1980: p. 450) explored the predictive effect of WM in language comprehension, including listening and reading, by emphasizing the individual differences in WM capacity. The measurements, memory span tasks, which they developed not only allowed them to reveal the predictor role of WM in language comprehension of their participants but also paved the way for other researchers in the field of the second or foreign language. Since the present research seeks to investigate the predictive role of WM in students' reading comprehension through C-DA, the next section will focus on the relationship between WM and reading ability, in particular the pioneering work of Daneman and Carpenter (1980: pp. 450-466) as well as important meta-analyses in the field (e.g. Daneman and Merikle, 1996: p. 422; Shin, 2020: p. 873).

### 2.5.1. Working Memory and Reading Comprehension

Language comprehension, including reading and listening, has been associated with WM, as students are expected to understand the semantic and syntactic relationships of the words, phrases, and sentences, as well as combine newly encountered knowledge with previously processed knowledge (Daneman & Merikle, 1996: p. 422). Assuming that good readers have more efficient skills of processing, there is a greater capacity to store partial products of the reading task (Harrington & Sawyer, 1992: p. 26). Daneman and Carpenter (1980: p. 450) specified that individual differences in language comprehension may be related to the differences

in WM capacity including the functions of process and storage. They stated that the basis for individual differences in WM is their functional capacity, where their limited capacity differs in maximum use rather than passive storage capacity; hence, those with functionally smaller storage capacity may encounter comprehension deficiencies, especially in the processes representing the words, phrases, and sentences in a coherence (Daneman & Merikle, 1996: p. 423). More specifically, the results of many studies have shown that WM is a good predictor of reading comprehension in both first language (L1) (e.g., Daneman & Carpenter, 1980: pp. 450-466) and L2 context (e.g., Harrington & Sawyer, 1992: p. 25) and it has also been supported by some meta-analysis (e.g., Daneman & Merikle, 1996: p. 422). Nevertheless, there are also other studies revealing the insignificant effect of WM on reading comprehension (e.g., Chun & Payne, 2004: p. 481).

In the reading span task created by Daneman and Carpenter (1980: p. 450), examinees are presented 60 separate sentences written in the cards one by one, asked to read them aloud, and remember the final word. In each set, the number of sentences varies from two to six sentences. At the end of each set, examinees are shown a blank card and they are required to remember the final words of the sentences in the order they are presented. Daneman and Carpenter (1980: p. 450), also developed a listening span task, the procedure of which is similar to the RST and they compared the predictor role of these WM tasks with the traditional word span and digit span tasks in different measures of reading comprehension. They concluded that reading and listening span tasks are good predictors of reading and listening comprehension in L1 context, but the results obtained from the traditional span tasks were not correlated with the language comprehension. Both the vital role of WM and the better predictive role of processing plus storage tasks rather than storage alone tasks in language comprehension are also supported by the meta-analysis by Daneman and Merikle (1996: p. 430). Inspired by the original version of RST, some researchers modified and developed different versions of the original RST (e.g., Alptekin & Erçetin, 2009: p. 627; Alptekin et al., 2014: p. 536; van den Noort et al., 2008: p. 35; Unsworth et al., 2009: p. 635; Waters & Caplan, 1996: p. 51). Some also integrate semantic judgement tasks into the processing part of WM

test (e.g., Shin et al., 2019: p. 320) while others benefit from grammaticality judgement tasks (e.g., Alptekin & Erçetin, 2010: p. 206).

Besides the importance of WM in the acquisition of L1 (e.g., Daneman, 1991: p. 445; Waters & Caplan, 1996: p. 51), its role in L2 has also been explored and is still being investigated (e.g., Alptekin & Erçetin, 2009: p. 627; 2010: p. 206; Harrington & Sawyer, 1992: p. 25; Leeser, 2007: p. 229; Miyake & Friedman, 1998: p. 339; Walter, 2004: p. 315). Within the field of L2 and more recently, Linck et al. (2014: p. 861) and Jeon and Yamashita (2014: p. 160) supported the positive correlation between WM and L2 reading comprehension in their meta-analyses, respectively. Further, Shin (2020: p. 1), in her meta-analysis, concluded that the relationship between WM and L2 reading comprehension is moderate. Among the methodology of the tasks; the scoring procedure, the language of the task, and the order of final word influence the relationship between two, and the reading task types also affect the extent to which WM involves in L2 reading comprehension. For example, Harrington and Sawyer (1992: p. 25) investigated the relationship between L2 RST scores and reading performance in TOEFL of advanced Japanese students learning English. Using digit, word, and reading span task consisting of simpler and shorter sentences than the original version in L2 and L1, their findings demonstrated a strong correlation between the RST scores and reading performance in TOEFL as well as a weak correlation between simple span tasks (word span and digit span) and reading comprehension in L2. Similarly, Shahnazari-Dorcheh and Adams (2014: p. 19) examined the relationship between the scores of WM (reading span) and reading test, revealing the effect of WM on reading ability of students with lower proficiency level. In one part of their study, Chang et al. (2019: pp. 458-472.) also found the predictor effect of WM in L2 reading comprehension of Chinese learners and positive correlations of WM with their writing and grammar.

In the Turkish Context, Alptekin and Erçetin (2009: p. 627; 2010: p. 206; 2011: p. 235) supported the findings on the positive relationship of WM with general reading comprehension and inferential comprehension in their studies. As an example, Alptekin and Erçetin (2009: pp. 627-639) also analyzed the relationship between WM and reading ability, emphasizing the role of proficiency level that

affects this relationship. They administered two WM measures (recognition and recall) and a reading comprehension test (literal and inferential) to the participants. The findings indicated a moderate significant correlation between composite WM scores (storage plus processing) and students' ability to infer in the text rather than their literal comprehension. Furthermore, the recognition-based task did not achieve to identify the individual variations in WM, unlike recall tasks. In another study of Alptekin and Erçetin (2010: p. 206), they investigated the relationship between L1 and L2 reading span tasks and L2 reading comprehension including the literal and inferential components. Their results indicated that storage capacity did not differ significantly by the language, namely L1 and L2, and there is a significant correlation only between L2 RST and inferential comprehension in L2. They also emphasized that the variations in the tasks and the procedure used in the studies may lead the different and inconsistent results related to WM and L2 reading ability (Alptekin & Erçetin, 2009: p. 629). In this regard, Leeser (2007: p. 253) also emphasized the effect of learners' prior content knowledge content (topic familiarity) in moderating the role of WM in reading comprehension.

Although most previous studies have revealed the significant role of WM in L2 reading, some researchers have reported inconsistent findings. To illustrate, Chun and Payne (2004: pp. 481-503) implemented the non-word repetition task and a version of RST by Daneman and Carpenter (1980: p. 450) to examine the individual differences in reading comprehension and vocabulary of 13 English students, who are in the second year of a German course. The findings did not yield a meaningful relationship of WM span with L2 reading comprehension as well as vocabulary acquisition. Leeser (2007: p. 253), in his study, suggested that when learners are familiar with the topics of the texts, they can benefit from higher WM in reading comprehension. In addition, Joh and Plakans (2017: pp. 8-9), based on their results, noted that the learners' prior knowledge influenced the WM's contribution to L2 reading comprehension and they concluded that learners need to have sufficient L2 and topic knowledge so that they can utilize their WM capacity effectively.

According to the findings of the studies mentioned above in the review of related literature, we can reach a two-way assumption. Firstly, we can conclude that

DA can be used as a diagnostic tool to determine not only learners' current level but also their potential level of development in the assessment procedure. In addition to increasing the performance of learners with the help of mediation provided, DA also promotes future development by integrating the instruction and assessment. Moreover, researchers have begun to implement interventionist C-DA which allows them to administer it to a large-scale group and get more comparable results among learners. C-DA is used to test learners' various skills such as reading comprehension, which is viewed as a vital skill for general language development. However, there is a lack of research on C-DA, so its applicability and effect need further investigation, especially in Turkey. Secondly, it can be assumed that WM has a positive correlation with L1 and L2 reading comprehension and most of the studies have reported its significant role in identifying individual differences in L2 reading comprehension. However, there are also some studies with contradictory findings of the relationship between WM and L2 reading ability. In response to the growing body of research in C-DA as well as the role of WM in L2 reading comprehension, the present research aims to examine the predictive role of WM in actual and mediated scores of L2 reading comprehension through a C-DA. It also seeks to compare the actual and mediated scores and to investigate how much students benefit from the mediation and how LPS differs among the students with identical actual scores.

# CHAPTER 3

## 3. METHODOLOGY

This section presents the methodology of the current thesis. Firstly, the research design along with the research questions is explained. Following this, the information about the participants and setting is described. Finally, the data collection instruments, procedure, and data analysis are also elaborated.

### 3.1. RESEARCH DESIGN

In this research following the interventionist C-DA framework, a quantitative approach was employed in the data collection and analysis because interventionist DA is based on the ZPD's quantitative interpretation (Poehner & Lantolf, 2005: p. 239). Like most of the C-DA studies, the present research utilized the cake format in which learners are immediately offered the mediational prompts when they have difficulty. Since this thesis is designed to reveal the predictive effect of WM on actual and mediated scores in the C-DA of L2 reading comprehension (i.e., whether there is a relationship between them), a correlational research design was used primarily. In addition, a comparative research design was also adopted to compare students' actual and mediated scores as well as the gain scores among the groups determined by the actual scores. Moreover, it was also aimed to present the results on LPS which groups students as high, mid, and low potential and sheds light on the planning of future instruction. Therefore, this research seeks to address the following research questions:

1. Does WM predict the actual scores in the C-DA of L2 reading comprehension?

2. Does WM predict the mediated scores in the C-DA of L2 reading comprehension?

3. Is there a significant difference between the actual and mediated scores in the C-DA of L2 reading comprehension?

4. Does the level of mediation use differ significantly among the students in the C-DA of L2 reading comprehension?

5. How does LPS differ among the students with the same actual score in the C-DA of L2 reading comprehension?

## 3.2. PARTICIPANTS AND SETTING

This research was carried out at a state university in Kocaeli, Turkey in the 2020-2021 academic year. The participants were 59 second-grade students in the department of ELT. They are within the age range of 19 to 47 with an average of 20.68 (*SD* = 4.05). Twenty-one of 59 students are male and 38 of them are female. All of them are native speakers of Turkish; most have been learning English for 10 years (*M* = 10.97, *SD* = 2.24) and they have had similar English learning experiences in formal education. Out of 59 students, 32 of them attended the English preparatory school before the four-year program in ELT department. Table 1 and Table 2 below provide the participants' background information.

**Table 1**

*Participants' Background Information Based on Age and Years of Learning English*

| | Number | Minimum | Maximum | Mean | SD |
|---|---|---|---|---|---|
| Age | 59 | 19 | 47 | 20.68 | 4.058 |
| Years of Learning English | 59 | 8 | 22 | 10.97 | 2.244 |

In the preparatory school, they had twenty-five class hours per week and they completed the eight-week English instruction including the courses based on grammar, vocabulary, reading, speaking, and listening. In addition, unlike other preparatory students, ELT students take courses from their own departments' lecturers, helping them get familiar with the department and the lecturers better. They have been learning EAP since they started the teacher education program in ELT department, namely for two years. They have some education courses where the medium of instruction is Turkish and they also take other courses of their department in English such as academic reading and writing as well as teaching approaches and methods in language teaching

**Table 2**

*Participants' Background Information Based on Gender and Attendance of Prep School*

| | | Prep School | | Total | Percentage |
|---|---|---|---|---|---|
| | | Yes | No | | |
| Gender | Female | 23 | 15 | 38 | 64.4 |
| | Male | 9 | 12 | 21 | 35.6 |
| Total | | 32 | 27 | 59 | 100 |
| Percentage | | 54.2 | 45.8 | 100 | |

Due to the nature of the task used for students' WM, we selected a more advanced group to ensure the comprehensibility of the scale based on the purposive sampling (Mackey & Gass, 2012: p. 185). Therefore, we included the students who are at the transition level as English proficiency, from EFL background, and currently studying EAP in the department of ELT. For the reading comprehension test, it was appropriate to include second-grade students who were supposed to be not novice readers. It was also aimed to include a group as homogeneous as possible, both demographically and academically. Therefore, students studying at the same grade at the same university were included to get more reliable scores and minimize the threats to internal validity. In the meantime, the easier access of the researcher to the group from the same institution was also taken into account in selecting the sample, and thus convenience sampling was also used (Mackey & Gass, 2012: p. 81). Further, it is difficult to handle the larger number of samples because all sessions were held online through one-to-one zoom meetings due to the pandemic period all around the world. The students have also participated in their courses online and they have had an intense schedule since the pandemic broke out in Turkey in March 2020. Hence, 59 students who have available time for the sessions volunteered for this research.

### 3.3. PROCEDURE

The data were collected in two different sessions, each lasting approximately 30-45 minutes over a nine-week period, with the exception of four weeks between October 2020 and January 2021. All of the sessions were held online through zoom

and the students were tested individually in one-on-one zoom meetings. The researcher planned the test administrations according to the students' available time by negotiating with them one by one taking into account their hunger, fatigue, and the intensity of online courses and exams. Prior to data collection, she informed the participants about the overall scope, length, and procedure of the research.

All of the participants were administered the WM task (RSPAN) at the first stage of this investigation. Since the form of the task was new for the students, the researcher spent 10 minutes explaining the procedure of the task and administering the trials to the participants before the real task. Most of the participants completed the RSPAN in 30 minutes or so. A few weeks later, they took the C-DA of reading comprehension as the second stage. The researcher has previous experience on face-to-face DA and C-DA, so she is also used to the mediation process. Before the tests started, she explained the necessary information and instructions as well as gave examples so that the students can understand the procedure more clearly. In addition to providing additional support for the task procedure, she tried to create a comfortable environment for students before the tests' administrations. Because the passages are intensive in meaning and vocabulary, the C-DA test took nearly 20-30 minutes. During both administrations, participants had the chance to ask what they found unclear and needed extra explanation. They also took the assessments seriously as the researcher also accompanied them. All WM and C-DA sessions were both audio and video-recorded with the consent of the participants to be used in data analysis; in case of any technical problem, the researcher also took notes based on the answers given while guiding them.

## 3.4. DATA COLLECTION INSTRUMENTS

### 3.4.1. Computerized Automated Reading Span Task (RSPAN)

To measure students' WM scores, a computerized automated reading span task (RSPAN) was administered. RSPAN, which is described by Unsworth et al. (2009: p. 635), requires not only storage (remembering the letters) but also processing (understanding the sentences) of information. According to the common view, since

complex span tasks involve the components of both storage and processing, they are better aligned with the WM capacity required in language processing tasks than simple span tasks (Waters & Caplan, 1996: p. 52). Further, traditional STM or WM measures, such as word span and digit span, are thought to be either unrelated or weakly associated with reading ability (Perfetti & Lesgold, 1977, as cited in Daneman & Carpenter, 1980: p. 450).

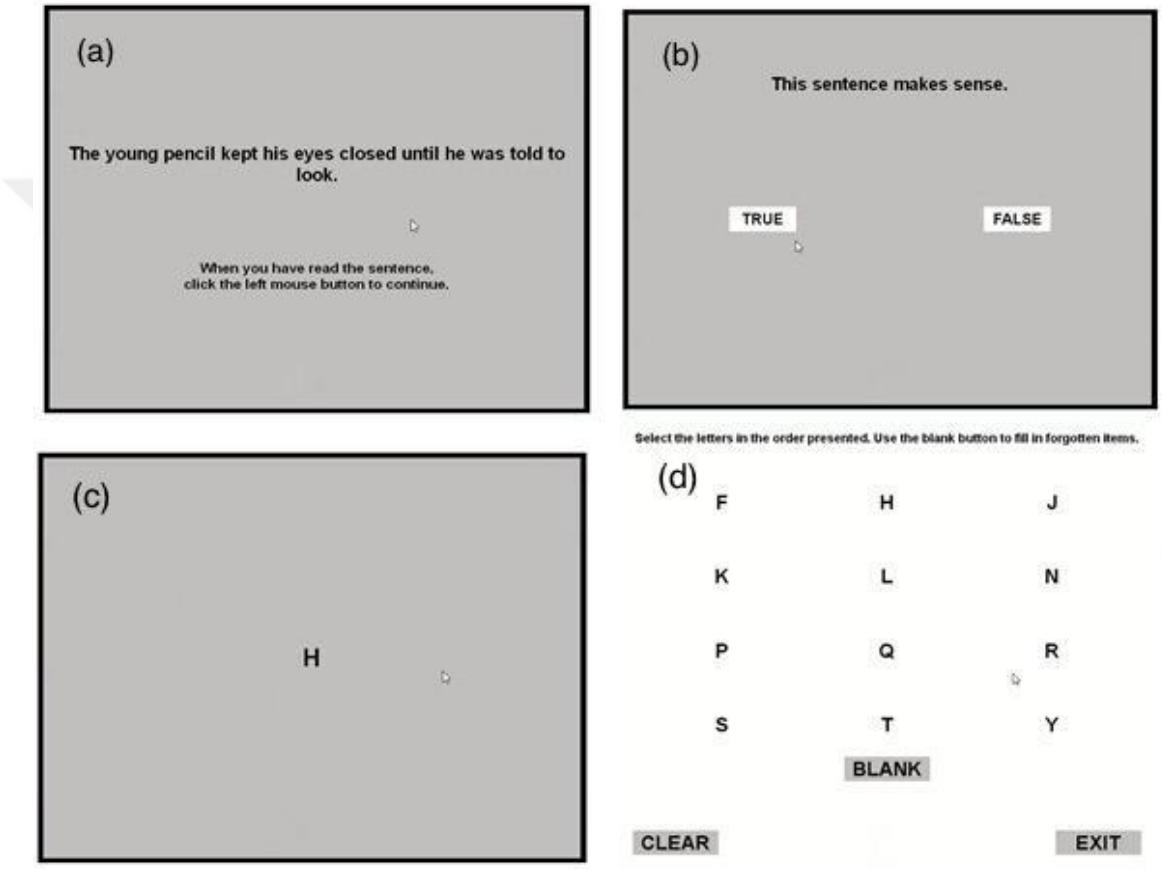
Unlike the traditional complex span tasks, automated tasks such as RSPAN are easily administered with the help of a computer and mouse and also provide scores automatically (Redick et al., 2012: p. 166). In their comprehensive study with 6000 college students, Redick et al. (2012: p. 166) reported that automated operation, symmetry, and reading span task are valid and reliable measurements, as evidenced by various analyses on the reliability and validity. Because some students do not have personal computers or mobile phones to download and run the required software (i.e., inquisit 6) for the task, it was administered one-on-one to the students through sharing the screen via zoom. Moreover, due to the pandemic period, students take their courses online; therefore, using this computerized task has become practical and easily accessible for both the researcher and the students.

For each level in the RSPAN task, participants are presented the sentences on the screen one by one (see Figure 5a). Appendix I presents some example sentences from the RSPAN. Immediately following each sentence, they are asked to make true/false judgments about the sentence in terms of making sense (see Figure 5b). After each sentence display, a letter is presented for later recall and the test taker can see the letter only for one second (see Figure 5c). After the series of sentences and letters are presented at each level, participants tried to remember the letters according to the order they appear on the screen by telling the letters aloud so that researcher could click them on a letter grid including 12 consonants (see Figure 5d). There are also some trials for the letter (storage task), sentence (processing task) and letter-sentence (processing and storage task) before the actual task begins. During these trials, the software calculates the average and individualized time of the examinee and while reading the sentences, if the examiner exceeds her average duration, the

task skips the true/false part and displays the next letter on the screen. Thus, the participants do not have extra time to memorize the letters while also reading the sentences. Figure 5 demonstrates the aforementioned steps of the RSPAN adapted from https://www.millisecond.com/download/library/rspan/.

**Figure 5**

*The Screenshots of the RSPAN Steps*




The RSPAN consists of 75 letters and 75 unrelated sentences with 10-15 words in length in the active voice. The test includes 15 levels ranging from three to seven sentences, and there are three sets for each three, four, five, six, and seven-sentence level. The number of sentences in the levels does not gradually increase; instead, for instance, the examinee can start at the six-sentence level and then move on to the three-sentence level. In this way, examinees cannot know how many sentences they will encounter and how many letters they need to memorize at the levels; they are required to memorize the letters until they see the letter grid on the screen to tell the

letters they remember for the researcher to click. In terms of the validity of the RSPAN, the researcher and an expert, who is also the students' lecturer, investigated the appropriateness of the sentences to the level of students and tested the task procedure. Furthermore, the findings regarding the reliability analysis of RSPAN showed that the reliability coefficient was .89, indicating an acceptable level.

### 3.4.2. Computerized Dynamic Assessment of Reading Comprehension

The C-DA was prepared by the researcher to measure students' actual and mediated reading comprehension scores. C-DA can be applied to a larger number of students than DA, and C-DA allows individuals to be re-evaluated when needed and their performance to be reported automatically (Poehner, 2008: p. 177). Furthermore, although interactionist DA has been used more in the past (e.g. Poehner & Lantolf, 2005), C-DA is becoming more and more popular among different researchers (e.g., Poehner & Lantolf, 2013: p. 323; Poehner et al., 2014: p. 337; Yang & Qian, 2017: p. 1; 2019: p. 1). The development of technology and the increase in the dependence on computers due to the pandemic spreading in the country have encouraged the researcher to implement C-DA to measure students' reading comprehension performance.

Participants' reading comprehension scores in English were measured through the use of short reading passages and associated multiple-choice questions. For the computerized reading comprehension test, passages and questions were adapted from the Graduate Record Examination (GRE) preparatory books and websites. GRE is required for the application of master's and doctorate programs of many institutions. It does not measure learners' knowledge of a specific content area; instead, it tries to predict how hard they need to work to understand the contents of the program they are selected. It includes the sections of analytical writing, verbal reasoning, and quantitative reasoning. In the verbal reasoning section, there are reading comprehension, text completion, and sentence equivalence questions; most of these are multiple-choice questions while some involve selecting a sentence from the passage and selecting one or more options from a list.

In GRE reading comprehension section, there are both short passages consisting of about 160 words and long passages with approximately 450 words. For the present project, short reading passages and three multiple-choice questions per text were selected. The passages include 150 words on average. The topics of the passages are literature, art, and the role of women. There are three multiple-choice questions for each text with a total of nine questions, and five options are provided for each question. The C-DA was prepared using the user-friendly Microsoft PowerPoint 2010. It is easy and practical to use in the preparation of such computerized tests. The prompts were placed in the options by adding hyperlinks. Once students choose the correct answer, they receive a prompt that the answer is correct and they can move on to the next question. When they chose a wrong option, the computer shows them that their answer is wrong and it presents the next mediation in a written form at the top of the slide. For this computerized reading comprehension test, a 64-page presentation was prepared. According to reliability analysis on actual and mediated scores in this C-DA, Cronbach's Alpha was found to be .834, which is an acceptable level of reliability. The criteria for selecting the test items and mediational phase will be explained in detail in the next sections.

#### 3.4.2.1. Selection of Test Items

Preparing an appropriate test that is suitable for students' ZPD is the most important factor in DA, thus the test should not be too simple or too difficult (Yang & Qian, 2017: p. 8). The main criteria for including GRE questions is that our participants study in the department of ELT, hence they may find easy any reading passage and answer the questions before the mediation process starts. Because the test would be in DA format including a procedure that gradually provides mediation, short GRE questions that are intense in terms of meaning and vocabulary despite being short were chosen to measure students' actual and mediated reading comprehension scores. Second, it is more appropriate to choose a test that students are not familiar with in order to ensure justice among students and to prepare a reliable test (Yang & Qian, 2019: p. 9). GRE was thought to be an unfamiliar and not a well-known test in our participants' age group as students are still at the undergraduate level in Turkey. Furthermore, the number of questions and lengths of

passages were deliberately chosen because longer texts and more questions may cause some problems with exhaustion and boredom. Students' capacity to focus on the computer screen was also taken into consideration because focusing on a computerized test that takes more than 30 minutes could be another problematic factor.

In order to choose appropriate passages and comprehension questions in accordance with the students' proficiency level and ZPD, the researcher consulted with the lecturer who currently teaches the students at the university to have general information about their level of proficiency and knowledge. Based on this information, she thoroughly reviewed GRE preparatory books, websites, and YouTube channels suggested by an American friend for GRE practice questions because the real GRE questions are not available. While selecting the passages and associated questions, the difficulty levels available on the website were taken into consideration to include the questions with similar difficulty. She set aside the passages and questions that would be included in the test in terms of length, difficulty, and following the same question type. Following that, she shared them with two colleagues to review in terms of the aforementioned aspects. Regarding their feedback, she eliminated the texts and questions that might not be appropriate for the content knowledge of the students.

Kozulin and Garb (2002: p. 119) suggested that we should include items testing cognitive abilities rather than prior knowledge of students for DA; that's why, the vocabulary questions were also eliminated because of this concern. The questions testing the skills of understanding explicit detail information, the purpose of the passage, and making inferences were included in the C-DA. Selected texts and questions were sent for a final review to an expert with a Ph.D. degree, who previously had experience in DA and C-DA, who is also the lecturer in a course of the students. All in all, it was decided to include three short reading passages, each containing three multiple-choice comprehension questions, for the C-DA. As a final step, the researcher simplified a few words in some questions considering that they can be problematic for students to understand.

#### 3.4.2.2. Preparing the Mediation

Mediation is any form of assistance offered in sequential order, from the most implicit to the explicit when learners have difficulty in their performance (Poehner, 2008: p. 38). During DA, learners are provided mediations based on the break-downs in their performance, by doing so, they can perform beyond their current abilities and mediation also creates instruction promoting students' developmental abilities (Poehner, 2009: p. 472). Mediation types are determined by learners' responsiveness during the assessment. The assessor provides both explicit and implicit mediation according to learners' ZPD to achieve higher mental functions (Lantolf and Poehner, 2010: p. 17-18). With the help of mediation such as hints, prompts or leading questions, not only the students' performances increase but also they develop strategies to transfer them in the new tasks. Poehner (2005: p. 257) also emphasized that ZPD not only helps students master a specific task but also provides them to transfer the understanding they developed to the other tasks.

Besides the importance of selecting appropriate questions, developing suitable and valid mediations is another essential issue in DA and the types of mediation differ according to the model used for the studies. Since the reading comprehension test was designed for a large-scale group through C-DA, the mediations should be based on the group's ZPD, not the individual students. Thus, pre-prepared prompts in the C-DA represent an interventionist model of DA. C-DA is not very flexible in providing mediations by the computer program, so it usually belongs to the interventionist DA (Yang & Qian, 2019: p. 4). Interventionist approaches provide results that are more easily comparable among students and it displays the required number of mediation during a specific task (Budoff, 1987, as cited in Poehner & Lantolf, 2013: p. 325). By this means, students' performances were scored based on the number of mediation used for each question.

The C-DA procedure includes silent reading and four levels of mediation for each question. The mediations are based on the principles of the regulatory scale of Aljaafreh and Lantolf (1994: p. 468); namely, they are "graduated" from implicit to

explicit and “contingent” which means that prompt is presented only when learners need it. Before organizing the mediations, it was determined what kind of mediation will be provided in each mediational step and the determined steps were arranged from implicit to explicit. All mediations offered to the students for each attempt are developed differently for each question according to the correct answers. While developing the mediations, the researcher also made use of the explanation parts of the answers in the books and websites where she adapted the question items. After the mediational moves were prepared and organized, two colleagues with an MA degree and an expert with a Ph.D. degree were asked to examine and review the mediational phases guiding the correct answers to have the validity of mediations in the C-DA process.

After students have finished reading a passage, they are asked a comprehension question with five options. They were asked to choose a correct answer from the given options. When their initial response is correct, they continue with the next question; however, if it is not, computer mediations are offered to them in an increasing sequence of explicitness. The computerized mediations for each question finish when they answer the question correctly and each student can get a maximum of four levels of mediation until they find the correct answer. Appendix II presents the sample mediation sequence in the C-DA prepared for this thesis and the Google drive link of the full test. The following mediational stages were predicated on while preparing the prompts for the questions:

***Mediation 1:*** If a student's initial answer is wrong, the first and most implicit mediation is presented, which is the same for all questions: “Your answer is incorrect. Can you please read the question again?” The main aim of the first mediation is to encourage students to review and reattempt the question.

***Mediation 2:*** If the first mediation does not lead students to the correct answer, the second prompt, which is more explicit than the first one, will be provided. It briefly explains what the question means, emphasizing the keywords in the question. To illustrate, if the question is related to the primary purpose of the text, the mediation guides students to find the main idea of the text or the author’s point in the text. As another example, if it is an inference question, the mediation highlights what the

student should make inferences about. In this way, students can try to find the relevant part where they can find the correct inference by focusing on the keywords presented in the mediation.

***Mediation 3:*** Third mediation will be offered when the student's third try is still incorrect. It leads the students to examine some underlined sentence(s) from the passage to find the correct answer, which narrows the search field to a few sentences. Poehner and Lantolf (2013: p. 331) also suggested that it was effective to draw student's attention to the part of the text containing the correct answer. Third mediation is also considered useful for those who do not understand where to search for the answer in the text even after highlighting the keywords of the questions.

***Mediation 4:*** The last mediation, which is the most explicit one, guides students to locate the most relevant part of the passage (a sentence, phrase, or keywords) where the student can find the correct answer. Thus, the last prompt narrows the search field further. It also provides some short explanations about the question and underlined part. That's why the last mediation also has an explanatory feature although it differs for each question. Last but not least, the last prompt provides more learning opportunities, especially for the students that do not fully understand the question because it points to the part in the text that is directly linked to the correct answer.

### 3.4.3. Scoring Scale

In the RSPAN, each correct response was given one point, making 75 the maximum score. All responses of the participants were recorded by the software and the total scores of the participants can be seen on the screen when the task ends. The task generates the absolute and total scores as well as total reading, accuracy, and speed errors at the end of the task. The absolute score shows "the sum of all trials in which all items were recalled in the correct serial order"; total score (It is referred to as "partial storage score" in the study of Redick et al. in 2012: p. 166) "is the sum of items recalled in the correct serial position, regardless of whether the entire trial was recalled correctly" (Redick et al., 2012: p. 166). In this research, to measure students' WM, the total scores were used regardless of whether they recalled the all letters in

each level because it was reported that total scores have higher internal consistency (Friedman & Miyake, 2005: p. 584; Redick et al., 2012: p. 169) and they are more associated with reading comprehension than absolute scores (Friedman & Miyake, 2005: pp. 581–590). Redick et al. (2012: p. 169) also suggested the total scores in terms of reliability and sensitivity to the individual differences in WM. For instance, based on the total score, a student gets four points provided that she remembers four letters out of five letters in the correct order in a five-sentence level.

In the C-DA of reading comprehension, the scoring procedure of Poehner and Lantolf (2013: pp. 331-332) was utilized, and based on this, two different numerical scores were obtained from the results of the C-DA. Each correct answer is awarded a score of four and the maximum score is 36. To measure students' actual scores (without mediation) which represent their independent performance without mediation, four points for the correct answer and zero point for the incorrect answer are provided. On the other hand, while grading students' mediated scores representing students' responsiveness to mediation during C-DA, the following grading procedure was used. If students find the correct answer without getting any mediation, they get four points. If they reach the correct answer on their second attempt for the same question, they get three points. As might be expected, if they answer correctly after receiving the second mediation- on their third try- they get two points and so on until only correct answer remains. Provided that they find the correct option following the last mediation, zero point is earned for that question because they have no other choice other than the correct one. All in all, actual scores are calculated taking into account the options students choose first whereas mediated scores are based on after which mediation the students find the correct answer. That is, their mediated scores decrease as the number of mediation used for the questions increases. Lastly, the gain scores representing the difference between scores with and without mediation were calculated as in the study of Poehner and Lantolf (2013: p. 334) and LPS was calculated with the formula which was developed by Kozulin and Garb (2002: p. 121) and adapted by Poehner and Lantolf (2013: p. 330) for the actual and mediated score, as seen in Figure 6.

**Figure 6**

*LPS Formula for Actual and Mediated Scores (Poehner & Lantolf, 2013: p. 330)*

$$LPS = \frac{(\text{S post-mediation} - \text{S actual})}{\text{Max S}} + \frac{\text{S post-mediation}}{\text{Max S}}$$

Abbreviated as follows:

$$LPS = \frac{2\,(\text{S post-mediation}) - \text{S actual}}{\text{Max S}}$$

## 3.5. DATA ANALYSIS

Data management and analysis were performed using Microsoft Excel 2010 and IBM SPSS 18.0. Firstly, students' RSPAN scores calculated by the software were recorded for each student and entered into SPSS as a quantitative variable. Secondly, total actual and mediated reading comprehension scores of the students were calculated according to the scoring procedure described in the previous section, that is, the order in which they chose the correct option. The researcher also benefited from the notes taken during the C-DA sessions and the video recording of the computer screen to grade students' responses. Two weighted numerical scores of students and the scores they got for each question (with and without mediation) were also inputted into SPSS as other variables. After the data were operationalized for analysis, the assumptions for the statistical analyses to be used were checked, which will be explained in the next sections.

### 3.5.1. Assumptions for Regression Analysis

To investigate the predictive effect of students' RSPAN scores on their actual (RQ1) and mediated scores (RQ2) in the C-DA of L2 reading comprehension, simple linear regression was performed. Regression analyzes were carried out separately on the actual and mediated scores, i.e. in different models. However, prior to these analyses, whether the model meets the assumptions for the simple linear regression

was examined. Initially, both the predictor variable (i.e., WM scores) and the outcome variables (i.e., actual and mediated scores) are continuous variables, which is necessary for the linear regression model. Second, the linearity assumption, which necessitates that the variables have a linear relationship, was evaluated with the scatter plot of the predictor variable against the outcome variables. Figure 7 indicates that there is a positive but weak linear relationship between WM and actual scores. Likewise, there is also a weakly positive linear relationship between WM and mediated scores, as seen in Figure 8. Although it is a little bit large, there still seems to be a linear scatter. These results were also confirmed by the Pearson's correlation coefficient of .12 for the WM-actual score relationship and .20 for the WM-mediated score relationship. According to Cohen's effect size, the relationship of WM with actual and mediated scores has a small effect size ($.1 < r < .3$) and corresponds to 1% and 4% variance respectively (Cohen, 1988: p. 79); it is therefore at applicable levels. Although there appears to be a number of outliers in the data, it was found that the maximum values of Cook's distance for actual (.113) and mediated scores (.088) were lower than 1, indicating that no influential outliers were detected, which supports the normality assumption (Tabachnick & Fidell, 2007: p. 75).

**Figure 7**

*Scatterplot of WM and Actual Scores*

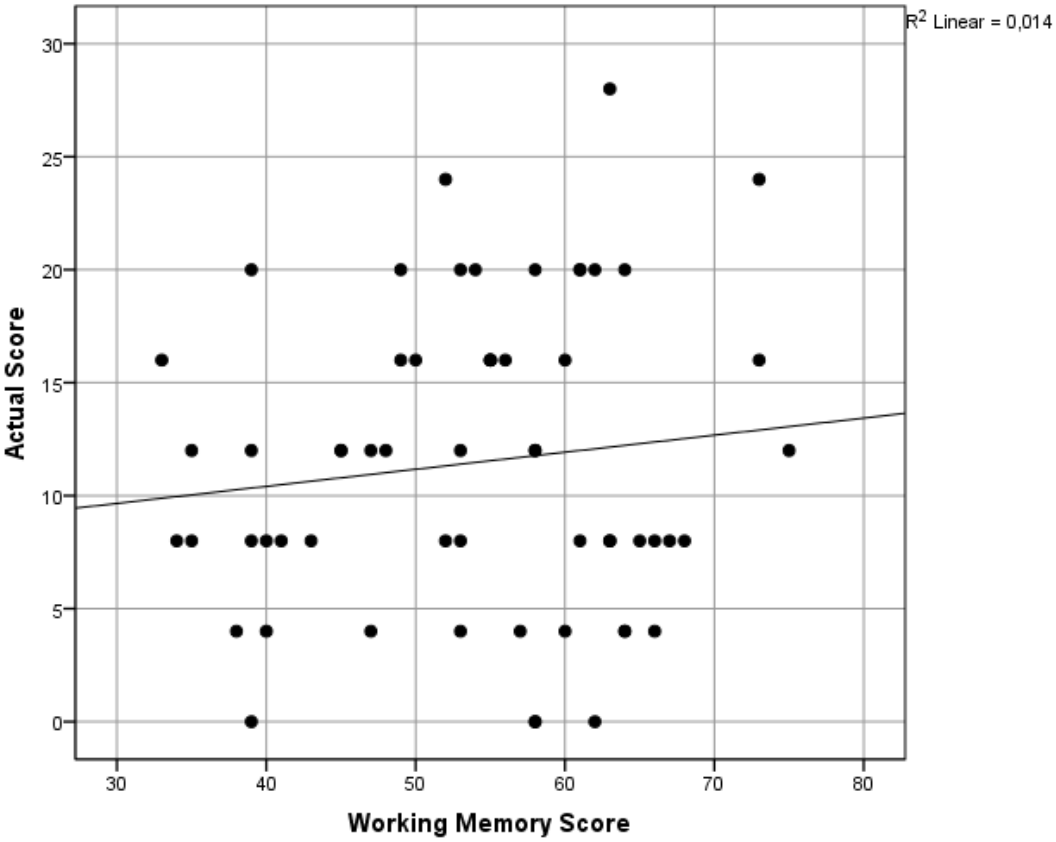
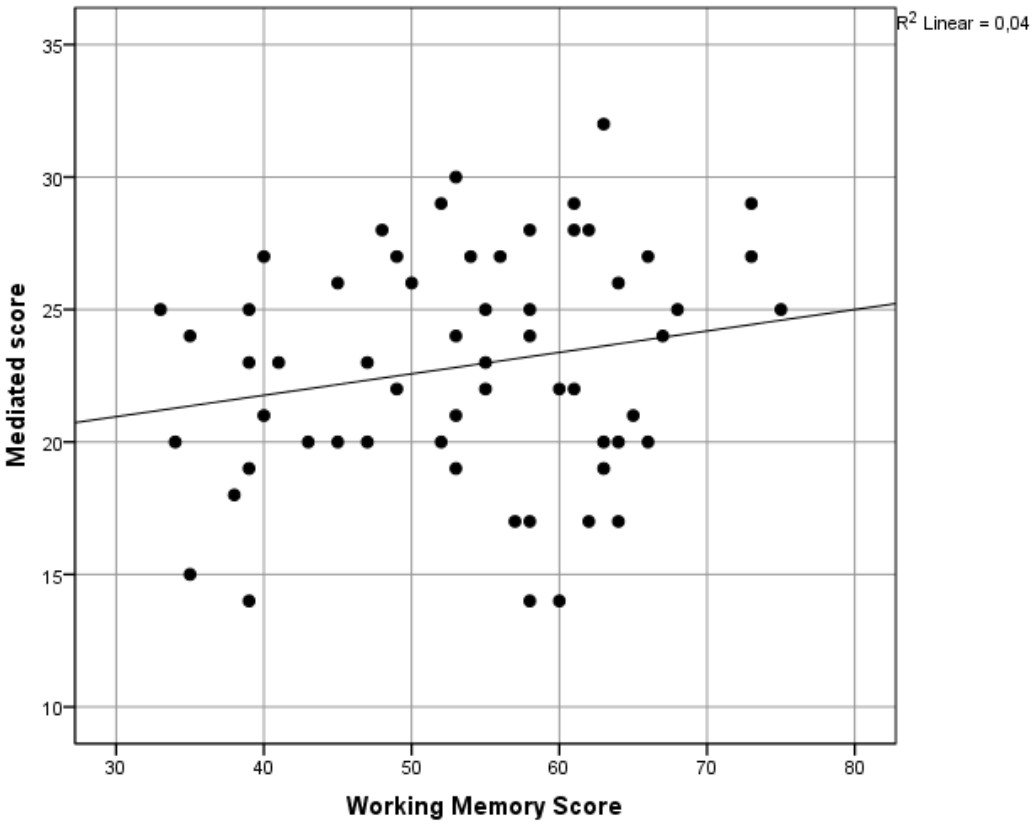


**Figure 8**

*Scatterplot of WM and Mediated Scores*




For the assumption of homoscedasticity that requires the constant error variance, the scatter plot of the predictor variable against the residuals (representing the errors that the regression model does not explain, i.e., the difference between the values predicted and observed) was visualized. Figure 9 (for actual scores) and Figure 10 (for mediated scores) demonstrate that the residuals are randomly scattered around the line without any clear pattern, which indicates that there is a constant variance, not heteroscedasticity (Larson-Hall, 2015: p. 259). The normality of the residuals assumption was also analyzed by the scatter plot of the residuals in both scores. The plots in Figure 9 and Figure 10 show that the errors are approximately normally distributed in both actual and mediated scores because the points are dispersed below and above the zero as well as to the graph's left and right (Larson-Hall, 2015: p. 259). In addition, since there is one independent variable as the predictor variable and there is not time-series data, other assumptions such as

multicollinearity, which might occur when more than one variable is highly correlated and autocorrelation ("independence of errors"), which may result from the time and distance of the cases (Tabachnick & Fidell, 2007: p. 88), were not included in the analysis. Taken together, a linear regression model was found to be appropriate for the data.

**Figure 9**

*Scatterplot of the WM against the Residuals (Actual Scores***)**

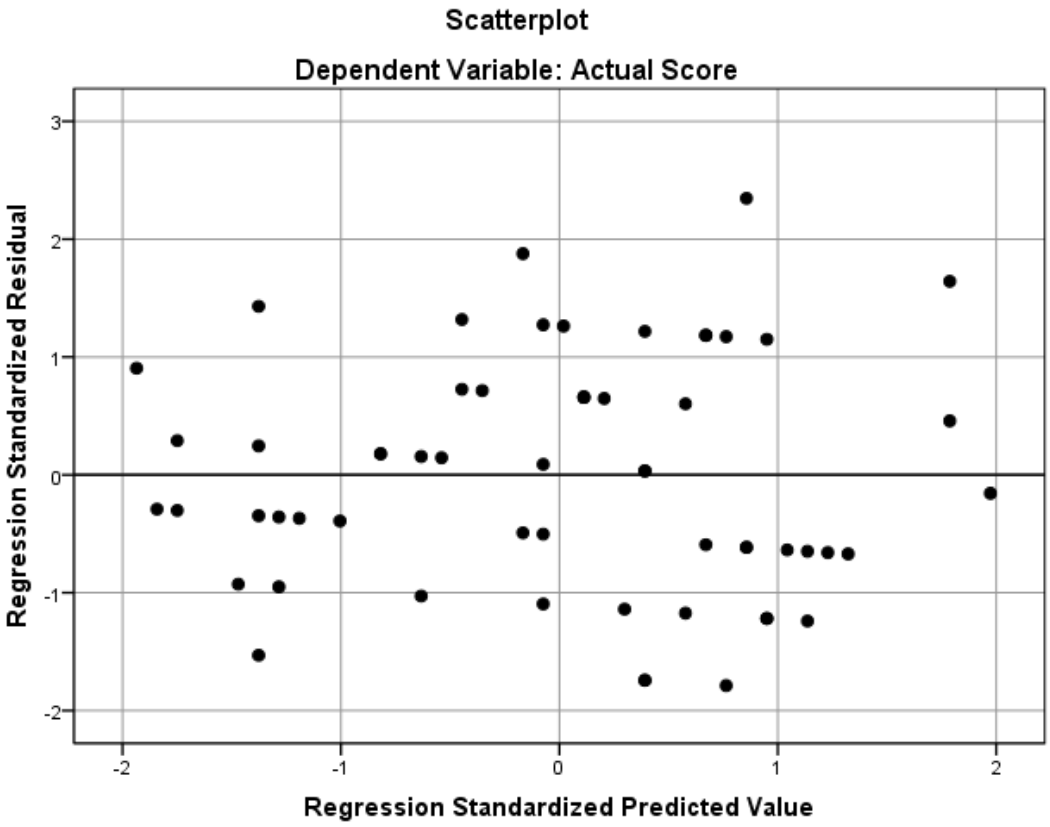
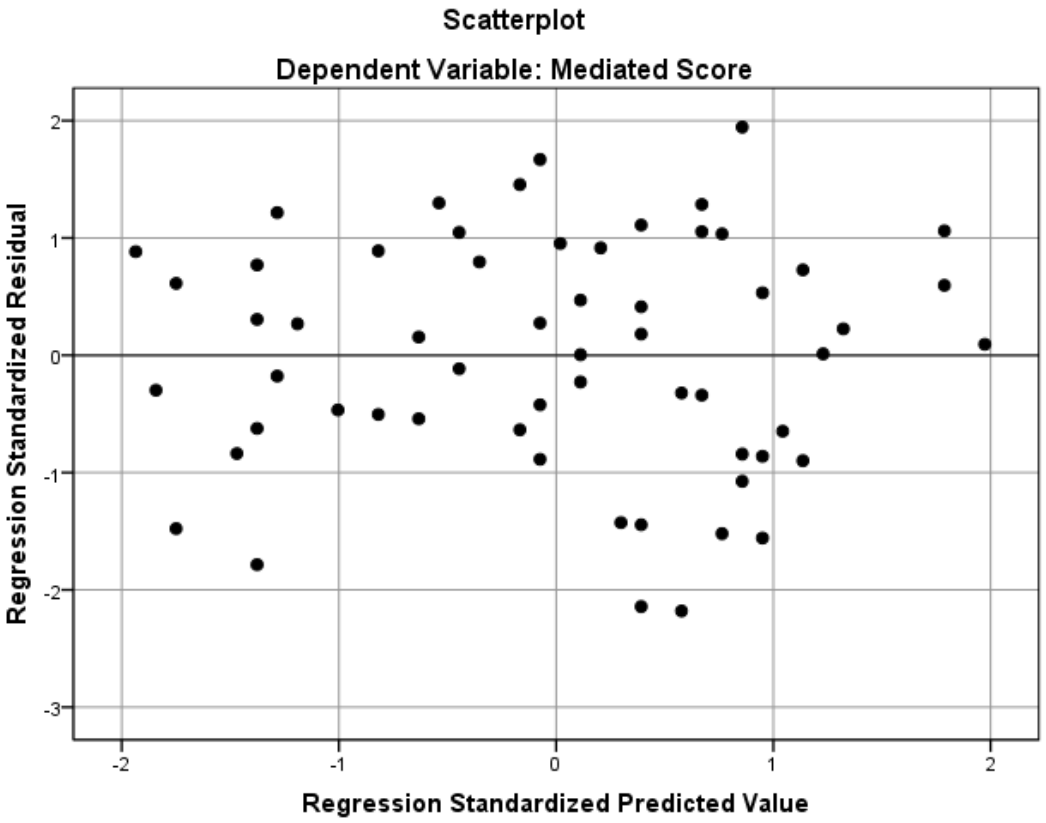


**Figure 10**

*Scatter Plot of the WM against the Residuals (Mediated Scores)*




### 3.5.2. Normality of the Data

The normality of the data obtained from all measurements was investigated using Kolmogorov-Smirnov test as this test is appropriate to use when the sample size is more than 50. According to the test of normality results, as seen in Table 3, there was a normal distribution in WM and mediated scores ($p > .05$); however, no normal distribution was observed in actual scores ($p < .05$). This might be explained by the fact that students' different reading proficiency levels may have affected the distribution of their actual scores, which was expected. The normal distribution of the mediated scores, on the other hand, indicated that the mediation process resulted in a more homogenized distribution of students' reading scores. Finally, there is no conflict regarding the normality of WM scores among the students. However, even if the scores deviate slightly from the normal distribution, test of normality results may

be significant (Field, 2013: p. 250). Therefore, descriptive statistics were also performed to elaborate the normality results through Skewness and Kurtosis of the data. As Table 3 demonstrates, Skewness and Kurtosis values in each calculated score are even stronger and within the acceptable range of + / -1 (Hair et al., 2010: p. 72). As a result, the scores were accepted as almost normally distributed and it was decided to analyze the obtained data with parametric tests.

**Table 3**

*Descriptive Statistics on the Normality of the Scores*

| | Kolmogorov-Smirnov | | | | | | |
|---|---|---|---|---|---|---|---|
| | Statistic | df | *p* | SD | SE | Skewness | Kurtosis |
| WM Score | .093 | 59 | .200* | 10.744 | 1.399 | -.222 | -.784 |
| Actual Score | .170 | 59 | .000 | 6.745 | .878 | .265 | -.621 |
| Mediated Score | .093 | 59 | .200* | 4.355 | .567 | -.215 | -.626 |

To investigate whether the actual and mediated scores obtained from the C-DA differ significantly (RQ3), Paired-Samples T-test was employed. To analyze the level of mediation use among the students (RQ4), they firstly were divided into three sub-groups based on their actual scores. After dividing the students into the groups, two outliers were detected and eliminated in the low group for this analysis. Because the number of students in the sub-groups was below 50, Shapiro-Wilk test was used to check the distribution of the gain scores within groups. Table 4 presents that the gain scores were normally distributed in each group, as evidenced by statistically non-significant Shapiro-Wilk values and Skewness and Kurtosis rates within -/+ 1 range (Hair et al., 2010: p. 72). According to the result of Levene's test, there is a homogeneous variance in the gain scores between groups ($p$ = .574). Therefore, one-way ANOVA was used because the data met the assumptions for this analysis. Finally, descriptive statistics were generated to examine how LPS differs among the students with the same actual score (RQ5).

**Table 4**

*Descriptive Statistics on the Normality of the Sub-groups*

| | Shapiro-Wilk | | | | | | |
|---|---|---|---|---|---|---|---|
| | Statistic | df | *p* | SD | SE | Skewness | Kurtosis |
| Low Achievers | .923 | 11 | .346 | 2.067 | .623 | -.751 | .837 |
| Mid Achievers | .961 | 34 | .258 | 2.830 | .485 | -.209 | -.253 |
| High Achievers | .961 | 12 | .793 | 2.491 | .719 | -.257 | -.956 |

# CHAPTER 4

## 4. RESULTS

This part of the thesis displays the findings obtained from the statistical analysis described in the previous chapter. First, the predictive effect of WM on actual and mediated reading comprehension scores in the C-DA was investigated. Second, the results of students' C-DA reading performance as well as the extent of mediation use and LPS were also reported and examined. The findings based on the quantitative analysis of two different task measures are presented in four main sections to address the research questions respectively. Initially, the descriptive statistics on the scores from all measures and the effect of WM on these scores are presented. It is followed by the comparative results on actual and mediated scores. Finally, the analyses of how the level of mediation use differs among the sub-groups (based on independent performance) and how LPS varies across the students with the same actual scores are displayed.

### 4.1. WORKING MEMORY AND L2 READING COMPREHENSION

#### 4.1.1. The Role of Working Memory in Actual Scores

To answer the first research question, simple linear regression was performed to examine how much variance in actual scores was explained by WM scores. In Table 5, the descriptive statistics demonstrate that the highest actual score was 28 (out of 36), the lowest was 0, with a range of 28 and a mean of 11.46 (*SD* = 6.745, *SE* = .878). On the other hand, the mean of WM scores was 53.80 (*SD* = 10.744, *SE* = .567). The highest WM score was 75 and the lowest was 33, with a range of 42. Table 6 shows that WM did not statistically significantly predict the actual L2 reading comprehension scores ($F(1,57) = 0.83$, $p > .05$). The B value was .75, demonstrating that actual scores increased by .75 for each point increase in WM scores. The $R^2$ value indicated that about 1.5% of the variance in the actual scores

was explained by WM, indicating a negligible effect of WM in actual scores, with a small effect size (Cohen, 1988: p. 79). In addition, the standardized Beta was found to be 0.120, meaning that actual scores increased by .120 standard deviations for each standard deviation increase in WM scores. Correlation analysis also revealed the positive but not significant relationship between WM and actual scores ($r$ = .120, $p$ >.05). All in all, WM was not a statistically significant predictor of reading comprehension in L2 without mediation.

**Table 5**

*Descriptive Statistics of All Scores*

| | N | Mean | SD | SE | Min | Max | Range |
|---|---|---|---|---|---|---|---|
| WM Score | 59 | 53.80 | 10.74 | 1.39 | 33 | 75 | 42 |
| Actual Score | 59 | 11.46 | 6.74 | .87 | 0 | 28 | 28 |
| Mediated Score | 59 | 22.88 | 4.35 | .56 | 14 | 32 | 18 |
| Gain Score | 59 | 11.42 | 4.28 | .55 | 2 | 23 | 21 |
| LPS | 59 | .9529 | .15 | .01 | .61 | 1.39 | .78 |

**Table 6**

*Linear Regression Output on WM and Actual Scores*

| | $\triangle R^2$ | *B* | *SE B* | Standardized $\beta$ | *t* value | *p* value |
|---|---|---|---|---|---|---|
| Constant | .014 | 7.396 | 4.527 | | 1.634 | .108 |
| Working Memory | | .075 | .083 | .120 | .915 | .364 |

### 4.1.2. The Role of Working Memory in Mediated Scores

In order to answer the second research question, a simple linear regression was again employed to analyze the relationship between WM and mediated scores in the C-DA of L2 reading comprehension. Descriptive statistics, as shown in Table 5, present that the mean of mediated scores was 22.88 (*SD* = 4.355, *SE* = .567). The lowest mediated score was 14 and the highest was 32 (out of 36), giving a range of 18. As Table 7 illustrates, simple linear regression demonstrated a statistically non-significant positive relationship between WM and mediated scores ($F$(1,57) = 2.36,

$p$ > .05). This finding was supported by the Pearson correlation coefficient of .200, representing a small effect size (Cohen, 1988: p. 79). The $R^2$ value revealed that 4% of the variance in the mediated scores can be accounted for by WM scores. The B value indicated that mediated scores increased by .08 with each point increase in WM scores while the standardized Beta value showed that mediated scores increase by .20 standard deviations for each standard deviation increase in WM scores. Overall, students' WM was not also a statistically significant predictor of their L2 reading comprehension performance with the mediation; however, the effect of WM on the mediated scores was higher than that on the actual scores.

**Table 7**

*Linear Regression Output on WM and Mediated Scores*

| | $\triangle R^2$ | *B* | *SE B* | Standardized $\beta$ | *t* value | *p* value |
|---|---|---|---|---|---|---|
| Constant | .040 | 18.530 | 2.885 | | 6.422 | .000 |
| Working Memory | | .081 | .053 | .200 | 1.537 | .130 |

## 4.2. ACTUAL AND MEDIATED SCORES

Descriptive statistics of the obtained scores (refer back to Table 5) illustrates that the average of the mediated scores ($M$ = 22.88) is higher than that of the actual scores ($M$ = 11.46). The standard deviation shows a decrease from the actual ($SD$ = 6.74) to the mediated score ($SD$ = 4.35), which is an indication that students show more homogenous performances in L2 reading when they receive mediation. Paired Samples t-test was used to compare the actual and mediated L2 reading comprehension scores in the C-DA. As Table 8 clearly demonstrates, a statistically significant difference between the actual and mediated scores was found ($t$(58) = -20.46, $p$ = .000). This finding revealed that the mediational prompts resulted in a noticeable improvement in the students' L2 reading comprehension in general during the C-DA. In terms of the relationship between the two scores obtained, a statistically significant positive relationship between the actual and mediated score was found as evidenced by a Pearson correlation coefficient of .784 ($p$ = .000), which is regarded as large effect size (Cohen, 1988: p. 80). Furthermore, this finding suggests that the

learners with higher actual scores tend to have higher mediated scores as well in the C-DA of L2 reading comprehension.

**Table 8**

*Paired-Samples T-test Results on Actual and Mediated Scores*

| | Paired Sample Statistics | | | Paired Differences | | | | |
|---|---|---|---|---|---|---|---|---|
| | N | Mean | SD | Mean | SD | *t* | df | *p* |
| Actual Score | 59 | 11.46 | 6.74 | | | | | |
| Mediated Score | 59 | 22.88 | 4.35 | -11.42 | 4.28 | -20.46 | 58 | .000 |

## 4.3. THE LEVEL OF MEDIATION USE: GAIN SCORES

To analyze the extent to which students benefit from the mediation in L2 reading comprehension through C-DA, gain scores, which represent the difference between the actual and mediated scores, were calculated for each student. To illustrate, a learner who receives an actual score of 20 and a mediated score of 30 is generated a gain score of 10. After that, students were divided into three sub-groups of high, low, and mid achievers based on their actual scores. To do this, first, the median of the actual scores was found to be 12. Second, about 30% of students on the right and left of the median were included in the mid achievers with a mean actual score of 11.29 ($n$ = 34); about 20% of the far right of the series was determined to be the high achievers with a mean actual score of 21.33 ($n$ = 12), and about 20% of the far left was included in the low achievers with a mean actual score of 2.55 ($n$ = 13).

**Table 9**

*One-way ANOVA Results on Gain Scores*

| | Sum of Squares | df | Mean Square | F | *p* | *η2* |
|---|---|---|---|---|---|---|
| Between Groups | 413.770 | 2 | 206.885 | 29.775 | .000 | .531 |
| Within Groups | 375.213 | 54 | 6.948 | | | |
| Total | 788.982 | 56 | | | | |

To compare the level of mediation use, i.e., gain scores, one-way ANOVA was employed. As Table 9 demonstrates, the results revealed that there was a statistically significant difference in terms of mediation use across the sub-groups (F(2,54) = 29.775, $p$ = .000, $\eta 2$ = .531). To examine which groups differed significantly in terms of gain scores and to compare the specific means of the groups, Hochberg's GT2 Post Hoc test was used because it is appropriate to use when the sample sizes differ from each other (Field, 2013: p. 555). The results (see Table 10) indicated that the mean gain score of the low achievers ($M$ = 14.45, $SD$ = 2.06) was significantly higher than the mean gain score of mid ($M$ = 11.59, $SD$ = 2.83) and high achievers ($M$ = 6.25, $SD$ = 2.49). This result suggests that the low group benefited most from the mediation among the other sub-groups in L2 reading comprehension through C-DA.

**Table 10**

*Gain Scores of Students with Low, Mid, and High Actual Scores*

| | Low Achievers (1) ($n$ =11) | | Mid Achievers (2) ($n$ =34) | | High Achievers (3) ($n$ =12) | | Post hoc |
|---|---|---|---|---|---|---|---|
| | M | SD | M | SD | M | SD | |
| Gain Score | 14.45 | 2.06 | 11.59 | 2.83 | 6.25 | 2.49 | 3<2<1 |

As we can see from Table 11, Post hoc analysis using Hochberg's GT2 also demonstrates that there was a statistically significant difference in the level of mediation use between the low and mid achievers ($p$ = .008) as well as the low and high achievers ($p$ = .000). Moreover, mid and high achiever groups also differed significantly in terms of gain scores ($p$ = .000) and mid achievers ($M$ = 11.59, $SD$ = 2.83) benefited from the mediation significantly more than the high achievers ($M$ = 6.25, $SD$ = 2.49). According to the correlation analysis, actual and gain scores were negatively and strongly correlated ($r$ = -.776, $p$ = .000). Figure 11 also visualizes that as the actual score increases, the mediated score also increases while the gain score decreases. Taken together, these findings suggest that the students with lower actual scores are more likely to benefit from the mediation than those with higher actual scores during the C-DA of L2 reading comprehension.

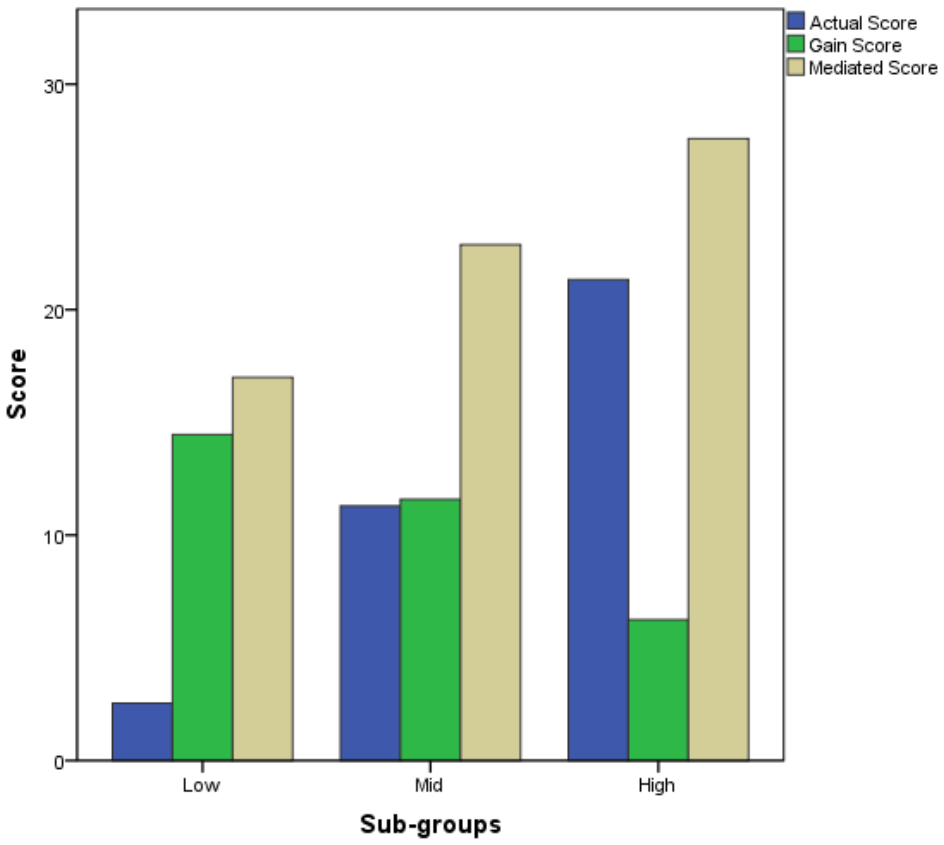
**Table 11**

*Multiple Comparisons on Gain Scores between Groups (Hochberg's GT2)*

| Actual Group (I) | Actual Group (J) | Mean Difference (I-J) | SE | *p* | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| Low | Mid | 2.866* | .914 | .008 | .62 | 5.12 |
| | High | 8.205* | 1.100 | .000 | 5.50 | 10.91 |
| Mid | High | 5.338* | .885 | .000 | 3.16 | 7.52 |

**Figure 11**

*Actual, Gain, and Mediated Scores of Students in Sub-groups*




## 4.4. LEARNING POTENTIAL SCORE (LPS)

Responding to the last research question investigating how LPS differs among the students with the same actual scores, students' LPSs were calculated according to the formula developed by Kozulin and Garb (2002: p. 121), based on the actual and mediated scores (refer back to Figure 2). LPS focuses on the gain between the actual

and mediated scores as well as the maximum score. In the study of Kozulin and Garb (2002: p. 121), the learners' LPSs ranged from 0.47 to 1.21 and they categorized the learners into three groups of high (LPS ≥ 1), low (LPS < .71), and intermediate learning potential group (.79 ≥ LPS ≤ .88). In this present research, the categorization of Kamrood et al. (2019: p. 18) was utilized because it has a clearer cut between groups. Following their study (Kamrood et al., 2019: p. 18), the students in this research were divided into three subgroups of low (LPS < .86), mid (.86 ≥ LPS ≥ 1.05), and high (LPS > 1.05) learning potential groups. The students in the mid learning potential group (*n* = 28) outnumbered the ones in the low (*n* = 15) and high groups (*n* = 16). Table 12 below illustrates the description of these groups based on LPS.

**Table 12**

*Descriptive Statistics on the Groups of Learning Potential*

| LPS range | Frequency | Percent | Mean | SD | SE |
|---|---|---|---|---|---|
| Low (LPS < 0.86) | 15 | 25.4 | .77 | .07 | .01 |
| Mid (0.86 ≥ LPS ≥ 1.05) | 28 | 47.5 | .95 | .04 | .00 |
| High (LPS > 1.05) | 16 | 27.1 | 1.12 | .11 | .02 |
| Total | 59 | 100 | .95 | .15 | .01 |

Fifteen students got the actual score of 8, which was the most frequent actual score, and it is followed by the actual score of 12 (*n* = 10), 16 (*n*= 9), and 20 (*n*= 9), respectively. As clearly seen in Table 13, the students with the same actual scores differentiate in their mediated scores and therefore LPS. Figure 12 also demonstrates the changing LPSs of the students with the same actual score of 8. To illustrate, students 2, 4, and 6 generated the same actual score of 8; however, students 2 and 4 had similar mediated scores, which are 20 and 19 in order, whereas student 6 received a higher mediated score (i.e., 24), indicating that she benefited more from the mediation. Furthermore, the LPS of student 6 was 1.11 while students 2 and 4 generated a LPS of .89 and .83 respectively. Therefore, they all fall into different learning potential groups despite their same independent performance in L2 reading comprehension.

**Table 13**

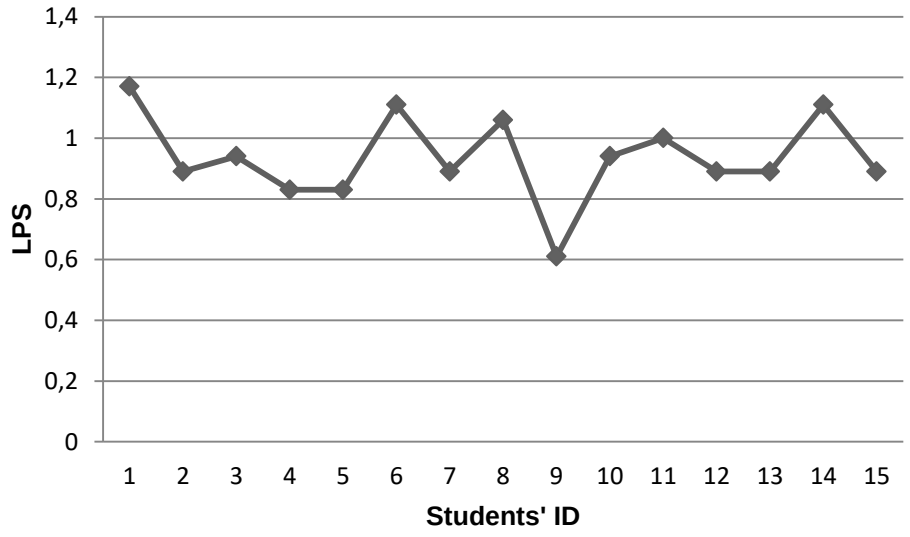
*The Scores of Students with the Same Actual Score*

| Student | Actual Score* | Mediated Score* | Gain Score | LPS |
|---|---|---|---|---|
| 1 | 8 | 25 | 17 | 1.17 |
| 2 | 8 | 20 | 12 | .89 |
| 3 | 8 | 21 | 13 | .94 |
| 4 | 8 | 19 | 11 | .83 |
| 5 | 8 | 19 | 11 | .83 |
| 6 | 8 | 24 | 16 | 1.11 |
| 7 | 8 | 20 | 12 | .89 |
| 8 | 8 | 23 | 15 | 1.06 |
| 9 | 8 | 15 | 7 | .61 |
| 10 | 8 | 21 | 13 | .94 |
| 11 | 8 | 22 | 14 | 1 |
| 12 | 8 | 20 | 12 | .89 |
| 13 | 8 | 20 | 12 | .89 |
| 14 | 8 | 24 | 16 | 1.11 |
| 15 | 8 | 20 | 12 | .89 |

*Note. * Maximum Score: 36*

**Figure 12**

*Different LPSs for Students with the Same Actual Score of 8*



As an extreme example, students 1 and 9 (see Table 13) got the same actual score of 8; the two students, however, performed quite differently through mediation. The performance of student 1 improved significantly under mediation, increasing her score by 17 points, while student 9 was only able to increase her score by 7 points. Consequently, student 1 with a LPS of 1.17 fell into the category of high learning potential; student 9, on the other hand, generated a low LPS of .61. Table 14 also displays how the scores varied among some students with identical actual scores. Several students, such as students 23 and 24, received higher actual scores than students 18 and 19 but were unable to improve their performance much through mediation, thus producing low-range LPSs. On the other hand, students 18 and 19 were involved in the high LPS group by improving their performance with the help of mediation. To summarize, the results revealed that the students who received the same or similar independent scores did not always have identical mediated scores in L2 reading comprehension. Instead, their openness to mediation, level of mediation use, and hence learning potential are more likely to differ.

**Table 14**

*The Scores of Students with Identical Scores*

| Student | Actual Score* | Mediated Score* | Gain Score | LPS |
|---|---|---|---|---|
| 16 | 12 | 24 | 12 | 1 |
| 17 | 12 | 20 | 8 | .78 |
| 18 | 12 | 25 | 13 | 1.06 |
| 19 | 12 | 28 | 16 | 1.22 |
| 20 | 16 | 22 | 6 | .78 |
| 21 | 16 | 25 | 9 | .94 |
| 22 | 16 | 27 | 11 | 1.06 |
| 23 | 20 | 22 | 2 | .67 |
| 24 | 20 | 26 | 6 | .89 |
| 25 | 20 | 30 | 10 | 1.11 |

*Note. * Maximum Score: 36*

# CHAPTER 5

## 5. DISCUSSION

This section focuses on interpreting and discussing the findings obtained from two measurements; a reading span task (RSPAN) and a C-DA of reading comprehension in L2. First, the results of students' WM, actual, mediated, gain, and learning potential scores are interpreted. Besides, the findings are discussed in light of the related literature on WM and L2 reading relationship as well as C-DA of L2 reading comprehension.

### 5.1. WORKING MEMORY AND L2 READING COMPREHENSION

This study primarily set out to investigate the predictive effect of WM on students' L2 reading comprehension performance with and without mediation in a C-DA. Although there are some contradictory findings on this topic, the correlation between WM and L2 reading comprehension has been reported in the literature in general; the findings of the current research, however, do not support previous common findings. The first research question sought to determine whether students' WM scores predict their actual (without mediation) L2 reading comprehension scores in the C-DA. The results revealed that WM was not a significant predictor of independent L2 reading comprehension performance and there was a weakly positive relationship between WM and reading comprehension in L2.

This result differs from the findings of some published studies (e.g., Alptekin & Erçetin, 2009: p. 627, 2010: p. 206; Harrington & Sawyer, 1992: p. 25; Jeon &Yamashita, 2014: p. 160; Linck et al., 2014: p. 861; Shin, 2020: p. 1; Walter, 2004: p. 315) indicating a moderate to strong relationship between WM and L2 reading comprehension. On the other side, this finding is in line with the study of Chun and Payne (2004: p. 481) that reported the insignificant impact of WM on reading comprehension in L2 and it also seems to consistent with the results of Joh and

Plakans (2017: pp. 8-9) concluding that the insignificant effect of WM on L2 reading of students with limited L2 and topic knowledge. A possible explanation for this result may be due to the familiarity of the students with the text topics in line with the findings of Leeser (2007: p. 253) and their prior knowledge or sufficient L2 knowledge as Joh and Plakans explained (2017: p. 8-9). Furthermore, as Shin (2020: p. 873) emphasizes, the role of WM in L2 reading may have been influenced by the language and scoring of the reading span task, as well as the form of reading comprehension task. In this context, Alptekin and Erçetin (2009: p. 629) reported that these inconsistent findings may be due to the differences in tasks and procedures used in studies.

On the other hand, the second research question aimed to find out the predictive effect of WM on students' L2 reading comprehension performance with mediation in the C-DA. The findings revealed that WM did not predict significantly the mediated L2 reading comprehension scores, as well. In other words, there was a weakly positive relationship between students' WM and their responsiveness to mediation during the C-DA. This finding suggests that the mediation process affects the contribution of WM to the reading comprehension in L2, albeit with a small effect size. Furthermore, the correlation of WM with the mediated scores is higher than with the actual scores, suggesting that students tend to utilize slightly higher WM under mediation. In reviewing the literature, no data was found on the association between WM and any type of DA performance including C-DA. Therefore, this research contributes to the literature by providing an insight into the relationship between WM and mediated reading performance in L2.

## 5.2. ACTUAL AND MEDIATED SCORES

The third research question investigates the comparison between the actual and mediated reading comprehension scores in the C-DA. The results showed a significant difference between students' performance with and without mediation, indicating that the mediation process contributed significantly to their L2 reading performance. This finding observed in the current research mirror that of earlier

studies (Barabadi, 2010: p. 1; Pishghadam et al., 2011: p. 1356; Pishghadam & Barabadi, 2012: p. 85-86; Poehner & Lantolf, 2013: p. 334; Poehner et al. 2014: p. 11; Teo, 2012; Yang & Qian, 2017: p. 10) reporting the significant difference between actual and mediated scores of learners. According to the results of standard deviations of these scores, students performed more homogenously when they receive mediation in the test similar to the finding of Poehner and Lantolf (2013: p. 334). Moreover, the correlation analysis confirms that there was a strong and positive correlation between the actual and mediated scores, revealing that students with higher independent scores had also higher scores under mediation. This finding corroborates the results of some studies (Kozulin & Garb, 2002: p. 120; Pishghadam & Barabadi, 2012: p. 86; Poehner et al., 2014: p. 11; Yang & Qian, 2017: p. 11) that revealed the significant relationship between these scores.

We should bear in mind that as Poehner and Lantolf (2013: p. 335) highlighted, mediated scores should not be interpreted as a direct reflection of learning; rather, they measure the degree to which students respond to the mediation, which is the process where learning takes place. Further, the goal of DA and C-DA is not to increase students' test scores but to learn about their ZPD, i.e., to determine their current and potential level of development (Poehner et al., 2014: pp. 3-4). Since actual or independent test scores represent only ZAD of the students, they do not provide any information about their emerging abilities. According to Vygotsky, in order to determine individuals' potential for learning, we should also take into account the ZPD. Moreover, students' responsiveness to mediation informs about not only their potential for future learning but also further instruction.

## 5.3. THE LEVEL OF MEDIATION USE: GAIN SCORES

According to Poehner and Lantolf (2013: p. 334), gain scores reflect "the change between the actual and mediated components of the tests". That is to say, gain scores indicate the degree to which students benefit from the mediation when there is a breakdown in their performance during the C-DA. Following this purpose, the fourth research question examined the level of mediation use among the students

with low, mid, and high actual L2 reading comprehension scores in the C-DA. The analysis of the results displayed that low achievers benefited significantly more from the mediation than students with mid and high actual scores. Moreover, mid achievers made use of mediation significantly higher than high achievers. Therefore, significant differences were found in terms of gain scores within each binary group across these three groups. These results agree with the study of Pishghadam et al (2011: p. 1356) who also separated students into two groups based on their independent scores. Their results showed that students' reading comprehension scores improved significantly in a dynamic test and low achievers benefited more from the mediational hints compared to high achievers.

In line with the findings of Poehner et al. (2014: p. 12) and Yang and Qian (2017: p. 11), the correlation analysis showed a strong negative correlation between the actual and gain scores. This finding supports the fact that students who did not perform well on their own took the advantage of mediation more than those who performed well independently. However, Poehner et al. (2014: p. 12) pointed that these results are due to the fact that students with high independent scores had less space for improvement under mediation. Therefore, these results on the difference between the scores need to be interpreted with caution. Furthermore, students' low scores in non-dynamic tests may result from the lack of learning opportunities or cultural differences as Pishghadam et al. (2011: pp. 1356-1357) suggested.

## 5.4. LEARNING POTENTIAL SCORE

The fifth question in this research aimed to examine how LPS differs among the students who receive identical actual reading comprehension scores in the C-DA. As stated in the previous sections, LPS reflects the difference between learners' independent and mediated performance, taking into account the maximum score and it allows practitioners to divide learners into groups with low, medium, and high learning potential. The findings showed that the students with similar or the same actual scores did not always have the same mediated scores, and therefore their LPSs also changed. In other words, students performing independently at the same level

may differ significantly in their responsiveness to mediation and learning potential. This result is in agreement with the findings of earlier studies (Poehner& Lantolf, 2013: p. 335; Poehner et al., 2014: p. 12; Yang & Qian, 2017: p. 11; Pishghadam & Barabadi, 2012: p. 88) that have examined the learning potential of learners through a dynamic approach. This finding also brings us to the same point as Vygotsky's (1956, as cited in Wertsch, 1985: p. 68) famous example, which demonstrated that two children of the same age level performed differently when mediated by examples or leading questions. In this regard, Vygotsky concluded that the children are very different in terms of their potential development despite their same actual developmental level (Vygotsky, 1956, as cited in Wertsch, 1985: p. 68). In a similar vein, this finding corroborates the ideas of Luria (1961: p. 7) who suggested that learners' performances differ significantly in a dynamic assessment in their ZPD although they might be similar according to statistical tests.

As supported by the overall findings of this research, actual scores provide information about only the current abilities of students; instead, C-DA informs about their current and potential development, i.e. ZPD, which would be more useful for diagnosis as well as future learning and teaching (Poehner et al., 2014: p. 13). In this regard, Poehner et al. (2014: p. 13) emphasized that when reporting the results of C-DA, we should focus on the actual and mediated scores together rather than on a single score; LPS also contributes to this interpretation by presenting the degree of improvement numerically. However, some points about the interpretation of the LPS are worth mentioning. To begin with, Kozulin and Garb (2002, as cited in Yang & Qian: p. 12) developed the formula to shed light on how much support students need to improve their skills in the future. Likewise, the researchers like Poehner et al. (2014: p. 10) used the term learning potential in the sense of "openness to mediation" in the context of L2 C-DA, which provides insight about further instruction based on the idea of Vygotsky's ZPD. Taken together, the scores of the students in this research gave information about their changing mediated performance despite their identical levels in a statistical approach. It should be noted once again that their changing performance indicates their different learning needs. Moreover, learning

potential scores and groups shed light on how much support they need and how they react to further teaching.

## CONCLUSION

The final chapter of this thesis provides a summary of the research. The findings are explained based on the research questions addressed. Moreover, the limitations of the research, pedagogical implications, and suggestions for further research are also explained.

## SUMMARY OF THE RESEARCH

This current research was designed mainly for a two-way purpose. It primarily set out (1) to investigate the predictive effect of WM on actual and (2) mediated L2 reading comprehension performance in an interventionist online C-DA and second to report and examine the C-DA results of students. Under the second purpose, it was aimed (3) to analyze whether there is a significant difference between students' actual and mediated performance; (4) to identify whether the level of mediation use (gain scores) differ significantly among the achievement groups based on actual performance, and (5) to investigate how LPS differentiates among the students with identical actual scores. The data were collected through two different tasks: a reading span task as a WM measure and a C-DA of reading comprehension test in L2 from 59 students who have studied EAP in the department of ELT. Students' actual, mediated, gain, and learning potential scores were calculated and analyzed quantitatively using Microsoft Excel 2010 and IBM SPSS 18.0. Accordingly, five research questions were addressed:

**Research Question 1: Does WM predict the actual scores in the C-DA of L2 reading comprehension?**

To analyze the predictive effect of WM on actual L2 reading comprehension performance, regression analysis was carried out. The results of this investigation indicate that WM scores through RSPAN did not predict significantly independent L2 reading comprehension scores in the C-DA; however, there was a weakly positive relationship between these scores with a small effect size. There are also some studies with contradictory findings although most of the studies support the moderate to strong positive association between WM and L2 reading ability. In this regard, the

current findings offer an inconsistent result compared to the common conclusion in the literature on this relationship.

As explained in the literature review and discussion section, some researchers pointed out the factors that affect the relationship between WM and reading comprehension. Inspired by their results, this result may be explained by the fact that students have different past language learning experiences, language assessment experiences, attitudes towards L2 reading and online assessments, whether they attended preparatory school, etc. Furthermore, general L2 knowledge (including vocabulary, grammar, etc.) inevitably affects L2 reading comprehension. In other words, if students do not have enough L2 knowledge to understand the texts, the involvement and contribution of WM to their language comprehension, therefore, may be limited. Since the participants are second-graders in the department of ELT, their language development is still ongoing, which also affects their reading performance. Finally, it is worth mentioning the students' views on the factors that may have influenced their results in this research. They noted that the fact that they have not taken such reading comprehension tests for a long time and took the test from the screen may also affect the results.

**Research Question 2: Does WM predict the mediated scores in the C-DA of L2 reading comprehension?**

The second question aimed to determine the predictive role of WM in mediated L2 reading comprehension scores in the C-DA. Regression analysis revealed that WM was not a significant predictor of mediated performance, i.e. responsiveness to mediation, during the C-DA of L2 reading comprehension. Correlation analysis also displayed again a weak positive association between these scores. WM also has a slightly higher effect on mediated scores than it does on actual scores, but with a small effect size. In other words, the interaction between the mediation and the students slightly increased the contribution of WM to L2 reading comprehension. It is difficult to explain this result, but it might be explained by the fact that the mediation, which is related to the correct answer, can activate the information that students read from the text and hold in their WM. This can also facilitate the

contribution of WM to their overall L2 reading performance during the C-DA. As mentioned earlier, this is the first research aimed at exploring the effect of WM on the C-DA performance revealing both current and potential levels of students. Therefore, the empirical findings in this research provide a new understanding of the relationship between WM and L2 reading comprehension with the involvement of mediation through a C-DA.

**Research Question 3: Is there a significant difference between the actual and mediated scores in the C-DA of L2 reading comprehension?**

To determine the effect of C-DA on students' L2 reading comprehension performance, the third research question sought to analyze whether the actual and mediated performance of students differed significantly. Paired Samples t-test results indicated that there was a statistically significant difference between the actual and mediated scores, suggesting that students generally were able to improve their reading performance with the help of mediational prompts during the C-DA. Moreover, the findings revealed a strong positive relationship between the actual and mediated scores, indicating that students who perform well independently also have high scores with mediation. The findings from this research make several contributions to the current literature. First, the findings add to the growing body of literature in the field of C-DA on L2 reading, where there is a lack of research, especially in Turkey. Second, the results have shown the diagnostic and significant effect of C-DA procedure on students' L2 reading, thus confirming previous findings of C-DA studies in different countries. Furthermore, these findings enhance our understanding of the applicability and advantages of an interventionist online C-DA in L2 context to a large number of students.

**Research Question 4: Does the level of mediation use differ significantly among the students in the C-DA of L2 reading comprehension?**

To investigate the degree to which students benefit from the mediation during the C-DA, gain scores were calculated for each student. Based on the actual scores, they were divided into low, mid, and high achievement groups and one-way ANOVA was conducted to analyze whether there is a significant difference among

these groups in terms of gain scores. The results revealed significant differences between all groups including low and mid, low and high, mid and high achievers. Furthermore, among the other groups, the students with low actual scores gained the most from the mediation. This was also confirmed by the correlation analysis, which indicated that as actual scores decreased, gain scores increased. These findings suggest that low achievers benefit more from a dynamic approach and that DA provides an opportunity for low achievers, in particular, to catch up with other students who perform well independently. This opportunity not only means that students increase their final scores, but also refers that with the help of mediation, they can continue to develop the knowledge and skills that they lack, even during the assessment.

**Research Question 5: How does LPS differ among the students with the same actual score in the C-DA of L2 reading comprehension?**

Students' LPSs were calculated using their actual and mediated scores as well as the maximum score of the test. They were also categorized into the low, mid, and high learning potential groups. To examine how and to what extent LPS varies among the students who received identical actual scores in L2 reading comprehension, their mediated and learning potential scores were reported and compared. The results indicated that even if students had the same independent reading score, their performance under mediation, i.e., their responsiveness to mediation, was generally different from each other, and thus their learning potential also differed. It is worth reminding once again that we use the term learning potential as openness to mediation in the field of L2 C-DA, rather than the potential in the areas such as intelligence or aptitude. LPS also allows students to be divided into different learning potential groups, which can offer important insights into how much support they should receive in the future learning process and how they can respond to this instruction. These findings lead us to the conclusion that statistical tests offer only students' actual developmental levels (i.e., ZAD), but do not show their potential levels (i.e., ZPD). Through a dynamic approach, we can diagnose their potential level with the help of mediational hints in an environment where the instruction is integrated into the assessment.

## LIMITATIONS OF THE RESEARCH

Although the results of this research make several contributions and new understanding to the current literature, there are some limitations to be considered when interpreting the results. Initially, although the validity and reliability checks have been performed in the research, the generalizability of the findings in wider contexts is subject to certain limitations because of the sample size. The second limitation is about the preparation of the computerized dynamic reading test and the mediational prompts for the questions. The computerized test was prepared by interviewing the students' lecturer at the university according to aspects such as their content knowledge and English language proficiency; the prompts were also arranged based on the guidelines of books and previous C-DA research. A pilot study involving dialogue between the mediator and students, on the other hand, would be more valuable in terms of checking the appropriateness of the questions and preparing the mediations more compatible with their ZPD. Finally, it should be noted that different mediators and mediational hints may result in different scores for students. As a result, the findings should be interpreted in light of the aforementioned limitations.

## PEDAGOGICAL IMPLICATIONS

The findings of this research have a number of important implications based on the analysis of students' reading span and C-DA reading comprehension performances in L2. First of all, DA gives information to educators about students' current as well as emerging abilities, i.e., their ZPD, as opposed to non-dynamic tests. In other words, scores generated by DA reveal their performance both with and without mediation. It is worth noting once again that students with the same ZAD do not always have the same ZPD, thus their performance may differ with mediation. That's why we recommend that students' developmental levels be interpreted with the help of a dynamic approach that encourages students to think and learn while being assessed, rather than the traditional tests. In this regard, LPS offers additional valuable information about individuals' proximal or future learning by dividing them into different learning potential groups. This enables program developers or

educators to create individualized plans for students and predict how students will respond to further instruction. Taken together, thanks to the diagnostic power of the dynamic approach, educators can benefit from all types of DA in both formal and informal settings to understand the whole picture of students' development (especially those they meet for the first time), to promote future development, and to prepare their future teaching more compatible with ZPD, as Vygotsky suggests.

Through a dynamic approach, educators can support their students in learning how to approach questions and using different strategies for different types of questions with the help of given prompts during the assessment. As students deal with the problematic parts in their performance with mediation in increasing order of explicitness, they can perform above what they can do independently, which integrates assessing and teaching. They can also internalize the knowledge and transfer it to future contexts, thus allowing them to continue their development. In this way, students can have more active roles in their own assessments and receive immediate feedback on their mistakes with the help of mediation, which can provide a kind of self-assessment. This is also an opportunity for both assessors and students to diagnose their strengths and weaknesses. Moreover, students, especially those who do not perform well independently, can be encouraged in terms of their increased scores at the end of the test, which allows an opportunity for them to catch up with high achievers. This, in turn, can reduce the stress and anxiety that students may experience during the examinations and make them feel comfortable because of the supportive environment with the mediator and mediational hints.

On the other hand, interventionist C-DA is more practical because it can be implemented to a large group of students and it provides comparable results through the numerical scores obtained. In terms of reliability, validity, or generalizability, interventionist DA is also more compatible with N-DA than interactionist DA, which increases the assessment's objectivity (Poehner, 2008: p. 45). As a solution to the workload of individual implementation of computerized dynamic tests, such as in this research, a software or a website can be developed so that a large number of students can take the test individually and simultaneously, and the scores can also be

generated automatically. This solution can be a great advantage for students, teachers, or language institutions at all levels, especially at a time when online and distance education are inevitable in our lives these days.

This research also provides detailed information about C-DA and C-DA studies from past to present and illustrates the preparation stages and the procedure; therefore, it might be a basis for future C-DA studies in Turkey and encourage practitioners in terms of the applicability and advantages of C-DA. However, it is also worth reminding that the test design used in this research, including the mediational prompts or the question type, is not a necessity to implement C-DA. Finally, the researcher believes that DA is an important part of language assessment literacy because it approaches language assessment from a different perspective than other types of assessment. For this reason, it is suggested that undergraduate and graduate programs in ELT departments incorporate the DA framework into their language assessment courses.

## SUGGESTIONS FOR FURTHER RESEARCH

In this research, current results regarding the relationship between WM and L2 reading comprehension contradict the general phenomenon in the literature. This enhances our understanding that WM may not have a significant impact on reading comprehension in all contexts, which requires further study. In future investigations, it might be possible to explore whether the contribution of WM to L2 reading comprehension increases as students' age and language proficiency increase in longitudinal or comparative research. It would also be interesting to explore the effects of training WM capacity of students on the contribution of WM to L2 reading comprehension in experimental research. Although some previous research highlighted the factors that affect WM's correlation with L2 reading comprehension, more information on these factors would help us to understand the role of different factors, or the combination of some factors, in this issue. For instance, prior to actual study, collecting data on language proficiency or prior knowledge such as vocabulary knowledge can also facilitate a more diagnostic analysis of students' reading

comprehension and the contribution of WM to reading. In addition, current research has provided a different perspective on the correlation of WM with L2 reading, showing that WM's contribution to L2 reading comprehension increased slightly with mediation in a dynamic approach. More research therefore is required to investigate the effect of WM on students' mediated reading performance.

On the other hand, as mentioned earlier, there is a general lack of research regarding the implementation of C-DA in L2 field, especially in Turkey. Previous C-DA studies including the current research revealed the advantages of the C-DA procedure over traditional static tests; however, further work using the computerized dynamic approach is still required to explain and confirm its effect on students' L2 performance. Another possible area of future research may be to investigate the long-term effect of C-DA on reading ability or any field within L2 development by integrating it into the curriculum and observing students' internalization and self-regulation processes over time. Furthermore, it is recommended that future studies be undertaken by incorporating more passages and questions into several C-DA sessions rather than a single session to get more comprehensive pictures of students' developmental levels. Different types of questions other than the multiple-choice format and transfer test items, which Poehner and Lantolf (2013: p. 324) suggested for the proof of learning, can also be included in C-DA. Finally, it is hoped that the findings of this research and the recommendations mentioned above will encourage further research on language assessment and development from a dynamic perspective.

## REFERENCES

### ***1. Books***


Ableeva, Rumia (2008). "The effects of dynamic assessment on L2 listening comprehension". In: Eds. James, P. Lantolf and Matthew, E. Poehner. Sociocultural Theory and the Teaching of Second Languages. London, UK: Equinox Pub, 57–86**.**

Afflerbach, Peter (2007). Understanding and using reading assessment. Newark, DE: International Reading Association.

Antón, Marta (2012). "Dynamic Assessment". In: Eds. Glenn Fulcher and Fred Davidson. The Routledge handbook of language testing. London, England: Routledge, 106-119.

Bachman, Lyle F. (1990). Fundamental considerations in language testing. Oxford: Oxford University Press.

Baddeley, Alan David, Graham J. Hitch (1974). "Working memory". In: Eds. G. Bower, The psychology of learning and motivation: Vol. 8. New York: Academic Press, 47–90.

Bodrova, Elena, & Deborah J. Leong (2003). "Learning and development of preschool children from the Vygotskian perspective". In: Eds. Alex Kozulin, Boris Gindis, Vladimir S. Ageyev, Suzanne M. Miller. Vygotsky's educational theory in cultural context. Cambridge: Cambridge University Press, 156-176.

Budoff, Milton (1987). "The validity of learning potential assessment". In: Ed. Carol S. Lidz. Dynamic assessment: An interactional approach to evaluating learning potential. New York: Guilford Press, 53–81.

Budoff, Milton (1968). Learning potential as a supplementary testing procedure. In: Ed. J. Hellmuth. Learning Disorders: Vol. 3. Seattle, WA: Special Child.

Brown, Ann L., Roberta A. Ferrara (1985). "Diagnosing zones of proximal development". In: Eds. James V. Wertsch. Culture, communication and cognition: Vygotskian perspectives. Cambridge: Cambridge University Press, 273–305.

Bruner, Jerome S. (1978). "The role of dialogue in language acquisition". In: Eds. Anne Sinclair, Robert J. Jarvelle, & Willem J. M. Levelt. The child's conception of language. New York: Springer- Verlag, 241-255.

Cazden, Courtney B. (2001). The language of teaching and learning. Portsmouth, NH: Heinemann.

Chaiklin, Seth (2003). “The zone of proximal development in Vygotsky‘s analysis of learning and instruction”. In: Eds. Alex Kozulin, Boris Gindis, Vladimir S. Ageyev, Suzanne M. Miller. Vygotsky’s educational theory and practice in cultural context. Cambridge: Cambridge University Press, 39-64.

Cohen, Jacob (1988). Statistical Power Analysis for the Behavioral Sciences. Hillsdale (NJ): Lawrence Erlbaum Associates.

Cole, Michael (1997). Cultural Psychology: A Once and Future Discipline. Cambridge: The Belknap Press of Harvard University.

Conway, Andrew R. A., Christopher Jarrold, Michael J. Kane, Akira Miyake, John N. Towse (2007). Variation in working memory. Oxford, UK: Oxford University Press.

Daniels, Harry (2016). Vygotsky and Pedagogy. NY: Routledge/Falmer.

Donato, Richard (1994). “Collective scaffolding in second language learning”. In: Eds. James P. Lantolf, Gabriela Appel. Vygotskian Approaches to Second Language Research. Norwood, NJ: Ablex, 33–59.

Ellis, Rod (2003). Task-based language learning and teaching. Oxford: Oxford University Press.

Feuerstein, Reuven, Yaacov Rand, John E. Rynders (1988). Don’t accept me as I am. New York: Plenum.

Field, Andy (2013). Discovering statistics using IBM SPSS statistics. London: Sage.

Fox, Janna (2017). “Using portfolios for assessment/alternative assessment”. In: Eds. Elana Shohamy, Iair G. Or, Stephen May. Language testing and assessment, Encyclopedia of language and education (3rd ed.). Cham: Springer, 135-148.

Guthke, Jürgen, Jens F. Beckmann (2000). “The learning test concept and its applications in practice”. In: Eds. Carol S. Lidz, Julian G. Elliott. Dynamic Assessment: Prevailing Models and Applications. Amsterdam: Elsevier, 17-69.

Hair, Joseph F. Jr., William C. Black, Barry J. Babin, Rolph E. Anderson (2010). Multivariate Data Analysis: A Global Perspective (7th ed.). London: Pearson Education.

Halliday, Michael A. K. (1978). Language as a Social Semiotics. London: Edward Arnold.

Hatch, Evelyn (1978). “Discourse analysis and second language acquisition”. In: Ed. E. Hatch. Second language acquisition: A book of readings. Rowley, Mass.: Newbury House, 401–435.

Haywood, H. Carl, Carol S. Lidz (2007). Dynamic Assessment in Practice: Clinical and Educational Applications. Cambridge: Cambridge University Press.

Hogan, Kathleen, Michael Pressley (Eds.) (1997). Scaffolding student learning: Instructional approaches & issues. Cambridge, MA: Brookline Books.

Karpov, Yuriy V. (2003). "Development Through the Lifespan". In: Eds. Alex Kozulin, Boris Gindis, Vladimir S. Ageyev, Suzanne M. Miller. Vygotsky's educational theory in cultural context. Cambridge: Cambridge University Press, 138-155.

Keane, Kevin J. (1987). "Assessing deaf children". In: Ed. Carol S. Lidz. Dynamic assessment: An interactional approach to evaluating learning potential. New York: Guilford, 360-376.

Kozulin, Alex (1998). Psychological tools: A sociocultural approach to education. Cambridge, MA: Harvard University Press.

Kozulin, Alex (2003). "Psychological Tools and Mediated Learning". In: Eds. Alex Kozulin, Boris Gindis, Vladimir S. Ageyev, Suzanne M. Miller. Vygotsky's educational theory in cultural context. Cambridge: Cambridge University Press, 15-38.

Lantolf, James P. (2000). "Introducing sociocultural theory". In: Eds. James P. Lantolf. Sociocultural Theory and Second Language Learning. Oxford: Oxford University Press, 1-26.

Lantolf, James P., Steven L. Thorne (2006). "Sociocultural theory and second language acquisition". In: Eds. B. VanPatten, J. Williams. Explaining SLA, 197-221.

Larson-Hall, Jenifer (2015). A guide to doing statistics in second language research using SPSS and R. New York: Routledge.

Leont'ev, Aleksei N. (1981). "The problem of activity in psychology". In: Eds. James V.Wertsch. The concept of activity in Soviet psychology Armonk, NY: M. E. Sharpe, 37–71.

Lidz, Carol S. (1987). Dynamic Assessment. New York, UK: Guilford Press.

Lidz, Carol S. (1991). Practitioner's Guide to Dynamic Assessment. New York, NY: Guilford Press.

Lidz, Carol S., Boris Gindis (2003). "Dynamic assessment of the evolving cognitive functions in children". In: Eds. Alex Kozulin, Boris Gindis, Vladimir S. Ageyev, Suzanne M. Miller. Vygotsky's educational theory in cultural context. Cambridge: Cambridge University Press, 99-116.

Lidz, Carol S., Julian G. Elliott (Eds.). (2000). Dynamic assessment: Prevailing models and applications. Amsterdam: Elsevier.

Mackey, Alison, Susan M. Gass (2012). Research methods in second language acquisition: A practical guide. Chichester: Wiley-Blackwell.

McCarthy, Sarah J., Susan McMahon (1992). “From convention to invention: Three approaches to peer interactions during writing”. In: Eds. Rachel Hertz-Lazarowitz, Norman Miller. Interaction in cooperative groups: The theoretical anatomy of group learning. Cambridge: Cambridge University Press, 17-35.

McNamara, Tim (2000). Language testing. Oxford: Oxford University Press.

Minick, Norris (1987). “Implications of Vygotsky’s theories for dynamic assessment”. In: Ed. Carol S. Lidz. Dynamic Assessment: An Interactive Approach to Evaluating Learning Potential. New York: Guilford, 116–140.

Miyake, Akira, Naomi P. Friedman (1998). “Individual differences in second language proficiency: Working memory as language aptitude”. In: Eds. Alice F. Healy, Lyle E. Bourne. Foreign Language Learning: Psycholinguistics Studies on Training and Retention, 339-361.

Newell, Allen (1973). “Production systems: Models of control structures”. In: Visual information processing. Oxford, England: Elsevier, 463-526.

Perfetti, Charles A., Alan M. Lesgold (1977). “Discourse comprehension and sources of individual differences”. In: Eds. Marcel Adam Just, Patricia A. Carpenter. Cognitive processes in comprehension. Hillsdale, N.J.: Lawrence Erlbaum Associates, 141–183.

Poehner, Matthew E. (2008). Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development (Vol. 9). Berlin: Springer Science & Business Media.

Poehner, M. E., Kristin J. Davin, James P. Lantolf (2017). “Dynamic Assessment”. In: Eds. Elana Shohamy, Iair G. Or, Stephen May. Language Testing and Assessment. Encyclopedia of Language and Education (3rd ed.). Cham: Springer, 243–256.

Rogoff, Barbara (1995). “Observing sociocultural activity on three planes: Participatory appropriation, guided participation, and apprenticeship”. In: Eds. James V. Wertsch, del Rio Pablo, Amelia Alvarez. Sociocultural studies of mind. Cambridge: Cambridge University Press, 139-164.

Samuels, Marilyn T. (2000). “Assessment of post-secondary students with learning difficulties: using dynamic assessment in a problem solving process”. In: Eds. Carol S. Lidz, Julian G. Elliott. Dynamic Assessment: Prevailing Models and Applications. Amsterdam: Elsevier, 521-542.

Schaffer, H. Rudolph (1996). Social development. Oxford: Blackwell Publishing.

Sternberg, Robert J., Elena L. Grigorenko. (2002). Dynamic Testing: The Nature and Measurement of Learning Potential. Cambridge: Cambridge University Press.

Tabachnick, Barbara G., Linda S. Fidell (2007). Using multivariate statistics (5th ed.) Boston, MA: Pearson.

Tharp, Roland G., Ronald Gallimore (1988). Rousing minds to life: Teaching, learning and schooling in social context. Cambridge: Cambridge University Press

Tzuriel, David (2001). Dynamic assessment of young children. New York, NY: Kluwer Academic/Plenum Publishers.

Van Lier, Leo (2004). The ecology and semiotics of language learning. Dordrecht: Kluwer Academic Press.

Vygotsky, Lev S. (1956). "Obuchenijia i umstvennoe razvitie v shkol'nom vozraste". In: Eds. Aleksei N. Leont'ev, Alexander R. Luria. Izbrannye psikhologicheskie issledovanija. Moscow: Izdatel'stov APN RSFSR, 438–452.

Vygotsky, Lev S. (1977). Play and its role in the mental development of the child. In M. Cole (Ed.), Soviet developmental psychology. White Plains, NY: M. E. Sharpe.

Vygotsky, Lev S. (1978). Mind in Society: The Development of Higher Psychological Processes. Cambridge, MA: Harvard University Press.

Vygotsky, Lev S. (1981). "The genesis of higher mental functions". In: Ed. James V. Wertsch. The concept of activity in Soviet psychology. Armonk, NY: M.E. Sharpe, 144-188.

Vygotsky, Lev S. (1986). Thought and Language. Newly Revised and Edited by Alex Kozulin. Cambridge, MA: MIT.

Vygotsky, Lev S. (1987). "Thinking and speech". In: Eds. Robert W. Rieber, Aaron S. Carton. The collected works of L. S.Vygotsky: (Vol.1) Problems of general psychology. New York: Plenum, 39-285.

Vygotsky, Lev S. (1994). "The development of academic concepts in school aged children". In: Eds. René van der Veer, Jaan Valsiner. The Vygotsky Reader. Oxford: Blackwell, 355–370.

Vygotsky, Lev S. (1998). "The Problem of Age". In Ed. Robert W. Rieber. The Collected Works of L. S. Vygotsky: (Vol. 5) Child Psychology. New York: Plenum.

Walsh, Steve (2013). Classroom discourse and teacher development. Edinburgh: Edinburgh University Press.

Wells, Gordon (1999). Dialogic inquiry: Towards a socio-cultural practice and theory of education. Cambridge: Cambridge University Press.

Wertsch, James V. (1985). Vygotsky and the social formation of mind. Cambridge, MA: Harvard University Press.

Wertsch, James V. (1993). Voices of the mind: A sociocultural approach to mediated action. Cambridge, MA: Harvard University Press.

***2. Articles, Papers, Other Printed Publications***

Ableeva, Rumia (2010). Dynamic assessment of listening comprehension in second language learning. Unpublished doctoral dissertation, The Pennsylvania State University, University Park, PA. Retrieved from https://etda.libraries.psu.edu/catalog/11063

Ahmadi, Alireza, Elyas Barabadi (2014). "Examining Iranian EFL learners' knowledge of grammar through a computerized dynamic test". Issues in Language Teaching, 3(2): 161-183.

Ajideh, Parviz, Nava Nourdad (2012). "The immediate and delayed effect of dynamic assessment on EFL reading ability". English Language Teaching, 12 (5): 141-151. doi: 10.5539/elt.v5n12p141

Aljaafreh, Ali, James P. Lantolf (1994). "Negative Feedback as Regulation and Second Language Learning in the Zone of Proximal Development". The Modern Language Journal, 78(4): 465-483. doi: 10.2307/328585

Alloway, Tracy Packiam, Ross G. Alloway (2010). "Investigating the predictive roles of working memory and IQ in academic attainment". Journal of Experimental Child Psychology, 106(1): 20–29. doi: 10.1016/j.jecp.2009.11.003

Alptekin, Cem, Gülcan Erçetin (2009). "Assessing the relationship of working memory to L2 reading: Does the nature of comprehension process and RST make a difference?". System, 37: 627–639.

Alptekin, Cem, Gülcan Erçetin (2010). "The role of L1 and L2 working memory in literal and inferential comprehension in L2 reading". Journal of Research in Reading, 33: 206–219.

Alptekin, Cem, Gülcan Erçetin (2011). "Effects of working memory capacity and content familiarity on literal and inferential comprehension in L2 reading". TESOL Quarterly, 45: 235–266.

Alptekin, Cem, Gülcan Erçetin, Oya Özdemir (2014). "Effects of variations in reading span task design on the relationship between working memory capacity and second language reading". The Modern Langauge Journal, 98(2): 536-552.

Antón, Marta (2009). "Dynamic Assessment of Advanced Second Language Learners". Foreign Language Annals, 42(3), 576–598. doi:10.1111/j.1944-9720.2009.01030.x

Baddeley, Alan D. (2000). “The episodic buffer: a new component of working memory?”. Trends in Cognitive Sciences, 4(11), 417–423.

Baddeley, Alan D. (2003).”Working memory and language: an overview”. Journal of Communication Disorders, 36(3), 189–208. doi: 10.1016/s0021-9924(03)000194

Barabadi, Elyas (2010). Designing computerized dynamic assessment of L2 reading comprehension of Iranian university students and its comparison with static test of L2 reading comprehension. Unpublished master’s thesis, Ferdowsi University of Mashhad, Mashhad, Iran.

Budoff, Milton, Martin Friedman (1964). “‘Learning potential’ as an assessment approach to the adolescent mentally retarded”. Journal of Consulting Psychology 28: 434–439.

Chang, Xin, Pei Wang, Meng-meng Cai, Mingzhu Wang (2019). “The predictive power of working memory on Chinese middle school students’ English reading comprehension”. Reading & Writing Quarterly, 35(5): 458-472.

Chun, Dorothy M., J. Scott Payne (2004). “What makes students click: Working memory and look-up behavior”. System, 32: 481-503.

Conway, Andrew R. A., Nelson Cowan, Michael F. Bunting, David J. Therriault, Scott R. B. Minkoff (2002). “A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence”. Intelligence, 30(2): 163–183. doi:10.1016/S0160-2896(01)00096-4

Cowan, Nelson (1988). “Evolving conceptions of memory storage, selective attention, and their mutual constraint withing the human information-processing system”. Psychological Bulletin, 104(2), 163-191.

Çalış, Selin (2018). The impact of the implementation of dynamic assessment on students‘ metasyntactic awareness. Unpublished master’s thesis, Çağ University, Mersin, Turkey.

Çetin-Köroğlu, Zeynep (2019). “Interventionist dynamic assessment’s effects on speaking skills testing: Case of ELT teacher candidates”. Advances in Language and Literary Studies, 10 (3), 23-31.

Daneman, Meredyth (1991). “Working memory as a predictor of verbal fluency”. Journal of Psycholinguistic research, 20(6), 445-464.

Daneman, Meredyth, Patricia A. Carpenter (1980). “Individual differences in working memory and reading”. Journal of Verbal Learning and Verbal Behaviour, 19: 450–466.

Daneman, Meredyth, Philip M. Merikle (1996). “Working memory and language comprehension: A meta-analysis”. Psychonomic Bulletin and Review, 3(4): 422-433.

d'Anglejan, Alison, Birgit Harley, Stan Shapson (1990). "Student evaluation in a multidimensional core French curriculum." Canadian Modern Language Review, 47(1): 106-124.

Davin, Kristin J. (2013). "Integration of dynamic assessment and instructional conversations to promote development and improve assessment in the language classroom". Language Teaching Research, 17(3), 303–322. doi: 10.1177/1362168813482934

Davoudi, Mohammad, Maryam Ataie-Tabar (2015). "The effect of computerized dynamic assessment of L2 writing on Iranian EFL learners' writing development". International Journal of Linguistics and Communication, 3(2): 176-186.

Dickson, Shirley V., David J. Chard, Deborah C. Simmons (1993). "An integrated reading/writing curriculum: A focus on scaffolding". LD Forum, 18(4): 12-16.

Ebadi, Saman, Abdulbaset Saeedian (2015). "The effects of computerized dynamic assessment on promoting at-risk advanced Iranian EFL students' reading skills". Issues in Language Teaching, 4(2): 1-26.

Ebadi, Saman, Abdulbaset Saeedian (2016). "Exploring transcendence in EFL learners' reading comprehension through computerized dynamic assessment". Iranian Journal of Language Teaching Research, 4(1): 27-45.

Ebadi, Saman, Abdulbaset Saeedian (2019). "Exploring L2 Learning Potential through Computerized Dynamic Assessment". Teaching English Language, 13(2): 51-78.

Ebadi, Saman, Hiwa Weisi, Hamed Monkaresi, Khosro Bahramlou (2018). "Exploring lexical inferencing as a vocabulary acquisition strategy through computerized dynamic assessment and static assessment". Computer Assisted Language Learning, 1–28. doi:10.1080/09588221.2018.1451344

Ebadi, Saman, Masoud Rahimi (2019). "Mediating EFL learners' academic writing skills in online dynamic assessment using Google Docs". Computer Assisted Language Learning, 32(5-6): 527-555. doi:10.1080/09588221.2018.152736

Engle, Randall W., Stephen W. Tuholski, James E. Laughlin, Andrew R. A. Conway (1999). "Working memory, short-term memory, and general fluid intelligence: A latent-variable approach". Journal of Experimental Psychology: General, 128(3): 309-331.

Estaji, Masoomeh, Abdulbaset Saeedian (2020). "Developing EFL learners' reading comprehension through computerized dynamic assessment". Reading Psychology, 41(4): 347-368.

Feuerstein, Reuven, Ya'acov Rand, Mildred B. Hoffman (1981). "The dynamic assessment of retarded performers: The learning potential assessment device,

theory, instruments and techniques". International Journal of Rehabilitation Research, 4(3): 465-466.

Friedman, Naomi P., Akira Miyake (2005). "Comparison of four scoring methods for the reading span test". Behavior Research Methods, 37: 581–590.

Grigorenko, Elena L. (2008). "Dynamic Assessment and Response to Intervention". Journal of Learning Disabilities, 42(2), 111–132. doi:10.1177/0022219408326207

Grigorenko, Elena L., Robert J. Sternberg (1998). "Dynamic testing". Psychological Bulletin, 124(1): 75-111.

Guterman, Eva (2010). "Toward dynamic assessment of reading: Applying metacognitive awareness guidance to enhance cognitive achievement". Journal of Research in Reading, 25(3), 283-298. doi: 10.1111/1467-9817.00176

Guthke, Jürgen (1982). "The learning test concept – An alternative to the traditional static intelligence test". The German Journal of Psychology 6(4), 306–324.

Güleryüz-Adamhasan, Betül (2019). Effects of dynamic assessment on EFL learners' agency. Unpublished doctoral dissertation, Çukurova University, Adana, Turkey.

Hargreaves, Andy, Lorna Earl, Michele Schmidt (2002). "Perspectives on alternative assessment reform". American Educational Research Journal, 30(1): 69–95.

Harrington, Michael, Mark Sawyer (1992). "L2 working memory capacity and L2 reading skill". Studies in Second Language Acquisition, 14(1): 25–38. doi:10.1017/S0272263100010457

Haywood, H. Carl, David Tzuriel (2002). "Applications and challenges in dynamic assessment". Peabody journal of education, 77(2): 40-63.

Hessamy, Gholamreza, Esmaeil Ghaderi (2014). "The role of dynamic assessment in the vocabulary learning of Iranian EFL learners". Procedia- Social and Behavioral Sciences, 98: 645-652. https://doi.org/10.1016/j.sbspro.2014.03.463

Hidri, Sahbi (2014). "Developing and evaluating a dynamic assessment of listening comprehension in an EFL context". Language Testing in Asia, 4(4), 1-19. https://doi.org/10.1186/2229-0443-4-4

Holmes, Joni, Susan E. Gathercole, Darren L. Dunning (2009). "Adaptive training leads to sustained enhancement of poor working memory in children". Developmental Science, 12(4): 9–15.

Huong, Le Pham Hoai (2007). "The more knowledgeable peer, target language use, and group participation". Canadian Modern Language Review, 64(2), 329–350.

Jacobs, Ellen L. (1998). “KIDTALK: a computerized language screening test”. Journal of Computing in Childhood Education, 9(2), 113–131.

Jacobs, Ellen L. (2001). “The effects of adding dynamic assessment components to a computerized preschool language screening test”. Communication Disorders Quarterly, 22(4), 217–226.

Jeon, Eun Hee, Junko Yamashita (2014). “L2 reading comprehension and its correlates: A meta-analysis”. Language Learning, 64(1): 160-212.

Joh, Jeongsoon, Lia Plakans (2017). “Working memory in L2 reading comprehension: The influence of prior knowledge”. System, 70(1): 107–120. doi:10.1016/j.system.2017.07.007

Just, Marcel Adam, Patricia A. Carpenter (1992). “A capacity theory of comprehension: individual differences in working memory”. Psychological review, 99(1): 122-149.

Kamrood, Ali M., Mohammad Davoudi, Saeed Ghaniabadi, Seyyed Mohammad Reza Amirian (2019). “Diagnosing L2 learners’ development through online computerized dynamic assessment”. Computer Assisted Language Learning, 1-30.

Kane, Michael J., David Z. Hambrick, Stephen W. Tuholski, Oliver Wilhelm, Tabitha W. Payne, Randall W. Engle (2004). “The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning”. Journal of Experimental Psychology: General, 133(2): 189-217.

Kar, Binod C., Udaya N. Dash, Jagannath P. Das, Jerry Carlson. (1993). “Two experiments on the dynamic assessment of planning”. Learning and Individual Differences, 5(1): 13–29.

Kayi-Aydar, Hayriye (2013). “Scaffolding language learning in an academic ESL classroom”. ELT journal, 67(3), 324-335.

Kır, Pınar (2020). The impact of interactionist dynamic assessment on EFL learners’ speaking skills and metacognitive awareness. Unpublished master’s thesis, Kocaeli University, Kocaeli, Turkey.

Klingner, Janette K., Sharon Vaughn (2000). “The helping behaviors of fifth graders while using collaborative strategic reading during ESL content classes”. TESOL Quarterly, 34(1): 69–98.

Kozulin, Alex, Erica Garb (2002). “Dynamic Assessment of EFL Text Comprehension”. School Psychology International, 23(1): 112–127. doi:10.1177/0143034302023001733

Lantolf, James P. (2009). “Dynamic assessment: The dialectic integration of instruction and assessment”. Language Teaching, 42(3), 355-368. doi:10.1017/S0261444808005569

Lantolf, James P., Aneta Pavlenko (1995). “Sociocultural theory and second language acquisition”. Annual Review of Applied Linguistics, 15: 108-124.

Lantolf, James P., Matthew E. Poehner (2004). “Dynamic assessment of L2 development: Bringing the past into the future”. Journal of Applied Linguistics, 1(1): 49-72.

Lantolf, James P., Matthew E. Poehner (2010). “Dynamic Assessment in the Classroom: Vygotskian Praxis for Second Language Development”. Language Teaching Research, 15(1): 11-33. doi: 10.1177/1362168810383328

Larkin, Martha J. (2001). “Providing support for student independence through scaffolded instruction”. Teaching exceptional children, 34(1), 30-34.

Leeser, Michael J. (2007). “Learner-based factors in L2 reading comprehension and processing grammatical form: Topic familiarity and working memory”. Language Learning, 57, 229–270.

Lee, Lina (2008). “Focus-on-form through collaborative scaffolding in expert-to-novice online interaction”. Language Learning & Technology, 12(3), 53–72.

Lehrer, Richard, Lee Shumow (1997). “Aligning the construction zones of parents and teachers for mathematics reform”. Cognition and Instruction, 15(1): 41-83.

Li, Mimi, Wei Zhu (2013). “Patterns of computer-mediated interaction in small writing groups using wikis”. Computer Assisted Language Learning, 26(1): 61–82. doi:10.1080/09588221.2011.631142

Lidz, Carol S. (1995). “Dynamic assessment and the legacy of LS Vygotsky”. School Psychology International, 16(2), 143-153.

Linck, Jared A., Peter Osthus, Joel T. Koeth, Michael F. Bunting (2014). “Working memory and second language comprehension and production: A meta-analysis”. Psychonomic Bulletin & Review, 21(4): 861–883. doi:10.3758/s13423-013-0565-2

Long, Michael H. (1981). “Input, interaction, and second language acquisition”. Annals of the New York academy of sciences, 379(1), 259–278).

Luria, Alexander R. (1961). “An objective approach to the study of the abnormal child”. Journal of the American Orthopsychiatric Association, 31, 1-16.

Ma, Qing (2017). “A multi-case study of university students’ language-learning experience mediated by mobile technologies: A socio-cultural perspective”. Computer Assisted Language Learning, 30(3–4), 183–203. doi:10.1080/09588221.2017.1301957

Malmeer, Elham, Masoud Zoghi (2014). “Dynamic assessment of grammar with different age groups”. Theory and Practice in Language Studies, 4(8): 1707-1713. doi:10.4304/tpls.4.8.1707-1713

McNamara, Danielle S. (2004). “SERT: Self-explanation reading training”. Discourse Processes, 38, 1–30. doi: 10.1207/s15326950dp3801_1

Mehri, Ehsan, Majid Amerian (2015). “Group dynamic assessment (G-DA): The case for the development of control over the past tense”. International Journal of Applied Linguistics and English Literature, 4(5): 11-20.

Miyake, Akira, Naomi P. Friedman, Michael J.Emerson, Alexander H.Witzki, Amy Howerter, Tor D. Wager (2000). “The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis”. Cognitive Psychology, 41: 49–100.

Naeini, Jila, Emily Duvall (2012). “Dynamic assessment and the impact on English language learners' reading comprehension performance”. Language Testing in Asia, 2(2): 22-41.

Nassaji Hossein, Merrill Swain (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. doi:10.1080/09658410008667135

Pica, Teresa (1994). “Research on negotiation: What does it reveal about second language learning conditions, processes, and outcomes?”. Language Learning, 44(3), 493–527. doi: 10.1111/j.1467-1770.1994.tb01115.x

Pica, Teresa, Richard Young, Catherine Doughty (1987). “The impact of interaction on comprehension”. TESOL Quarterly, 21(4): 737-758.

Pishghadam, Reza, Elyas Barabadi (2012). “Constructing and Validating Computerized Assessment of L2 Reading Comprehension”. Iranian Journal of Applied Linguistics, 15(1): 73-95.

Pishghadam, Reza, Elyas Barabadi, Ali M. Kamrood (2011). “The Differing Effect of Computerized Dynamic Assessment of L2 Reading Comprehension on High and Low Achievers”. Journal of Language Teaching & Research, 2(6): 1353-1358.

Poehner, Matthew E. (2005). Dynamic assessment of oral proficiency among advanced L2 learners of French. Unpublished doctoral dissertation, The Pennsylvania, State University, Pennsylvania, The USA.

Poehner, Matthew E. (2007). “Beyond the test: L2 dynamic assessment and the transcendence of mediated learning”. The Modern Language Journal, 91(3), 323–340. doi:10.1111/j.1540-4781.2007.00583.x

Poehner, Matthew E. (2009). “Group Dynamic Assessment: Mediation for the L2 Classroom”. TESOL Quarterly, 43(3), 471–491. doi:10.1002/j.1545-7249.2009.tb00245.x

Poehner, Matthew E., James P. Lantolf (2005). “Dynamic assessment in the language classroom”. Language Teaching Research, 9(3): 233–265. doi:10.1191/1362168805lr166oa

Poehner, Matthew E., James P. Lantolf (2013). “Bringing the ZPD into the equation: Capturing L2 development during computerized dynamic assessment (C-DA)”. Language Teaching Research, 17(3): 323-342.

Poehner, Matthew E., Jie Zhang, Xiaofei Lu (2014). “Computerized dynamic assessment (C-DA): Diagnosing L2 development according to learner responsiveness to mediation”. Language Testing, 32(3): 337–357. doi:10.1177/0265532214560390

Rea-Dickins, Pauline, Sheena Gardner (2000). “Snares and silver bullets: Disentangling the construct of formative assessment”. Language Testing, 17(2): 215-243.

Redick, Thomas S., James M. Broadway, Matt E. Meier, Princy S. Kuriakose, Nash Unsworth, Michael J. Kane, Randall W. Engle (2012). “Measuring Working Memory Capacity with Automated Complex Span Tasks”. European Journal of Psychological Assessment, 28(3): 164–171. doi:10.1027/1015-5759/a000123

Rosenshine, Barak, Carla Meister (1992). “The use of scaffolds for teaching higher-level cognitive strategies”. Educational leadership, 49(7): 26-33.

Savage, Robert, Natalie Lavers, Vanitha Pillay (2007). “Working memory and reading difficulties: What we know and what we don’t know about the relationship”. Educational psychology review, 19(2): 185-221.

Schneider, Elke, Leonore Ganschow (2000). “Dynamic assessment and instructional strategies for learners who struggle to learn a foreign language”. Dyslexia, 6: 72–82.

Shabani, Karim (2012, February). “Computerized dynamic assessment: An interventionist procedure to assess L2 reading”. In 6th National and 3rd International Conference of E-Learning and E-Teaching. Tehran, Iran: IEEE: 15-24.

Shabani, Karim, Mohamad Khatib, Saman Ebadi (2010). “Vygotsky's Zone of Proximal Development: Instructional Implications and Teachers' Professional Development”. English language teaching, 3(4): 237-248.

Shah, Priti, Akira Miyake (1996). “The separability of working memory resources for spatial thinking and language processing: An individual differences approach”. Journal of Experimental Psychology: General, 125(1): 4-27.

Shahnazari-Dorcheh, Mohammadtaghi, Rebecca Adams (2014). “The relationship between working memory and L2 reading comprehension”. Applied Research on English Language, 3(2): 19–34.

Sharafi, Mohammad, Sedigheh A. Sardareh (2016). "The Effect of Dynamic Assessment on Elementary EFL Students' L2 Grammar Learning". Journal of Applied Linguistics and Language Research, 3(3): 102-120.

Shin, Jihye, Vedran Dronjic, Boonjoo Park (2019). "The interplay between working memory and background knowledge in L2 reading comprehension". TESOL Quarterly, 53: 320–347.

Shin, Jihye. (2020). "A meta-analysis of the relationship between working memory and second language reading comprehension: Does task type matter?". Applied Psycholinguistics, 41(4), 873-900.

Siwathaworn, Prathana, Jirada Wudthayagorn (2018). "The impact of dynamic assessment on tertiary EFL students' speaking skills". The Asian Journal of Applied Linguistics, 5 (1): 142- 155.

Stevenson, Claire E., Bergwerff, Catharina E. Bergwerff, Willem J. Heiser, Wilma C. M. Resing (2014). "Working memory and dynamic measures of analogical reasoning as predictors of children's math and reading achievement". Infant and Child Development, 23(1): 51-66.

Summers, Robert (2008). Dynamic assessment: towards a model of dialogic engagement. Unpublished doctoral dissertation, University of South Florida, Florida, the USA.

Swain, Merrill (2000). "The output hypothesis and beyond: Mediating acquisition through collaborative dialogue". Sociocultural theory and second language learning, 97, 114.

Swanson, H. Lee (2006). "Working Memory and Dynamic Testing in Children with Learning Disabilities". Working Memory and Education, 125–156. doi:10.1016/b978-012554465-8/50007-7

Şentürk, Yeliz (2019). Dynamic assessment of secondary school EFL learners' tense preferences in writing activities. Unpublished master's thesis, Kocaeli University, Kocaeli, Turkey.

Teo, Adeline (2012). "Promoting EFL students' inferential reading skills through computerized dynamic assessment". Language Learning & Technology, 16(3), 10-20.

Thouësny, Sylvie (2010, June). "Assessing second language learners' written texts: An interventionist and interactionist approach to dynamic assessment". In EdMedia+ Innovate Learning. Toronto, Canada: Association for the Advancement of Computing in Education (AACE): 3517-3522. Retrieved from: http://www.editlib.org/p/34803

Turner, Marilyn L., Randall W. Engle (1989). "Is working memory capacity task dependent?". Journal of Memory and Language, 28(2): 127-154.

Tzuriel, David, Adina Shamir (2002). “The effects of mediation in computer assisted dynamic assessment”. Journal of Computer Assisted Learning, 18(1): 21–32. doi:10.1046/j.0266-4909.2001.00204.x

Ulu, Büşra (2020). Dynamic assessment of the use of metadiscourse markers in EFL writing. Unpublished master’s thesis, Kocaeli University, Kocaeli, Turkey.

Unsworth, Nash, Thomas S. Redick, Richard P. Heitz, James M. Broadway, Randall W. Engle (2009). “Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage”. Memory, 17(6): 635-654.

van den Noort, Maurits, Peggy Bosch, Marco Haverkort, Kenneth Hugdahl1. (2008). “A standard computerized version of the reading span test in different languages”. European Journal of Psychological Assessment, 24(1): 35-42.

Veen, Chiel van der, Marjolein Dobber, Bert van Oers (2016). “Implementing dynamic assessment of vocabulary development as a trialogical learning process: A practice of teacher support in primary education schools”. Language Assessment Quarterly, 13(4): 329-340. doi: 10.1080/15434303.2016.1235577

Walter, Catherine (2004). “Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory”. Applied Linguistics, 25, 315–339.

Wang, Peihui (2015). “The effect of dynamic assessment on the listening skills of lower-intermediate EFL learners in Chinese technical college: A pilot study”. Journal of Language Teaching and Research, 6(6), 1269-1279. doi:10.17507/jltr.0606.14

Waters, Gloria S., David Caplan (1996). “The Measurement of Verbal Working Memory Capacity and Its Relation to Reading Comprehension”. The Quarterly Journal of Experimental Psychology Section A, 49(1): 51–79. doi:10.1080/713755607

Weismer, Susan E., J. Bruce Tomblin, Xuyang Zhang, Paula Buckwalter, Jan Gaura Chynoweth, Maura Jones (2000). “Nonword repetition performance in school-age children with and without language impairment”. Journal of Speech, Language, and Hearing Research, 43(4): 865–878.

Wiedl, Karl H. (2003). “Dynamic testing: A comprehensive model and current fields of application”. Journal of Cognitive Education and Psychology, 3(2), 93-119.

Wood, David, Jerome S. Bruner, Gail Ross (1976). “The role of tutoring in problem solving”. Journal of Child Psychology and Psychiatry, 17: 89–100.

Yang, Yanfeng, David D. Qian (2017). “Assessing English reading comprehension by Chinese EFL learners in computerized dynamic assessment”. Language Testing in Asia, 7(1): 1-15.

Yang, Yanfeng, David D. Qian (2019). "Promoting L2 English learners' reading proficiency through computerized dynamic assessment". Computer Assisted Language Learning, 33(5-6): 628-652. doi:10.1080/09588221.2019.158588

Yılmaz-Yakışık, Burçak (2012). Dynamic assessment of ELT students' speaking skills. Unpublished doctoral dissertation, Gazi University, Ankara, Turkey.

Yılmaz-Yakışık, Burçak, Abdulvahit Çakır (2017). "Dynamic assessment of prospective English teachers' speaking skill". European Journal of English Language Teaching, 2(1): 22-53.

Yuill, Nicola, Jane Oakhill, Alan Parkin (1989). "Working memory, comprehension ability and the resolution of text anomaly". British Journal of Psychology, 80: 351-361.

Zadeh, Hasan H. (2018). "The Impact of Dynamic Assessment on EFL Learners' Vocabulary Learning". Annals of Language and Literature, 2(3), 1-7.

***3. Electronic Resources***

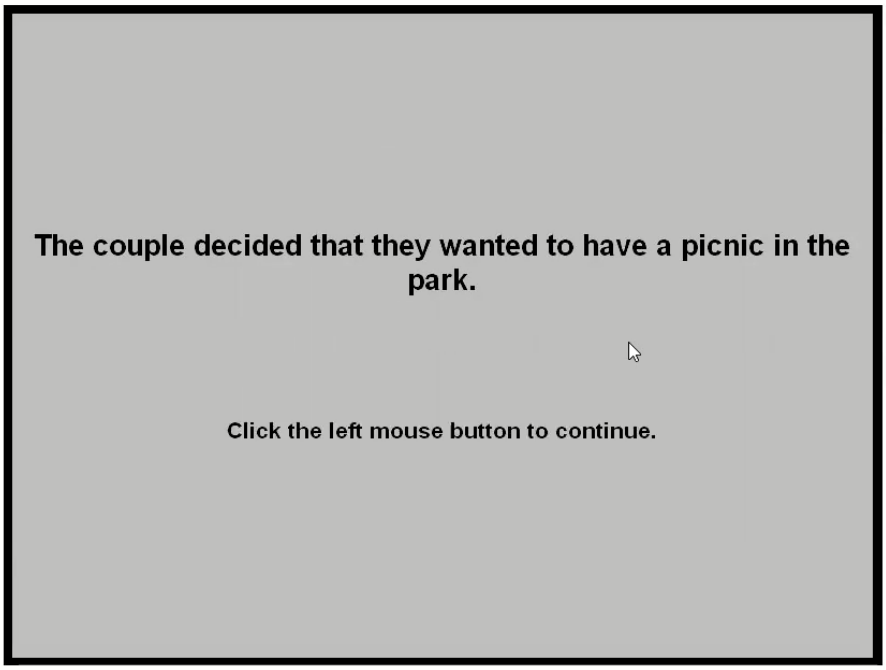
Lisina, Maya I. (1986). Problemy ontogeneza obscheniya [Problems of the ontogenesis of communication]. Moscow: Pedagogika. Retrieved from http://psychlib.ru/mgppu/LOl-1997/LOl-0172.htm#$p17

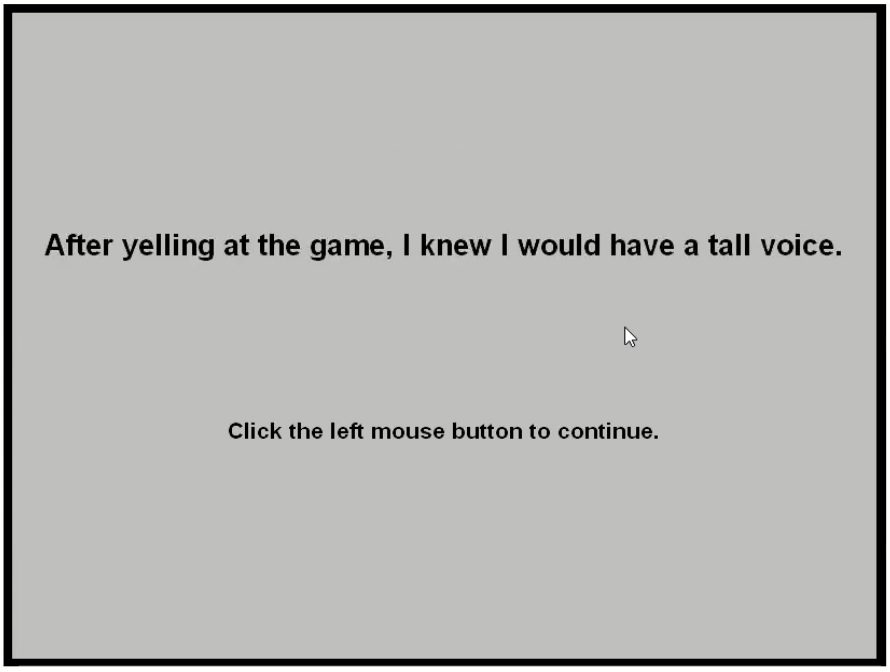
McLeod, Saul (2007). Lev Vygotsky. Simply Psychology. Retrieved from http://www.simplypsychology.org/vygotsky.html

## APPENDICES

**APPENDIX I.** Screenshots of Some Example Sentences from the RSPAN

The RSPAN is available at the link below:
https://www.millisecond.com/download/library/rspan/



The boys knew they would have to hurry to make it to the apple on time.

Click the left mouse button to continue.

They were worried that all of their luggage would not fit in the car.

Click the left mouse button to continue.

**APPENDIX II.** An Example of the Mediation Sequence of a Question from the C-DA of Reading Comprehension

The full version of the C-DA is available at the Google drive link below:

https://drive.google.com/file/d/1cByXC0MkD7PD_WHKiVmprx1CSxYeGrav/view?usp=sharing

Many critics of Emily Brontë's novel Wuthering Heights see its second part as a counterpoint that comments on, if it does not reverse, the first part, where a romantic reading receives more confirmation. Seeing the two parts as a whole is encouraged by the novel's sophisticated structure, revealed in its complex use of narrators and time shifts. Granted that the presence of these elements need not argue for an authorial awareness of novelistic construction comparable to that of Henry James, their presence does encourage attempts to unify the novel's heterogeneous parts. However, any interpretation that seeks to unify all of the novel's diverse elements is bound to be somewhat unconvincing. This is not because such an interpretation necessarily stiffens into a thesis (although rigidity in any interpretation of this or of any novel is always a danger), but because Wuthering Heights has recalcitrant elements of undeniable power that, ultimately, resist inclusion in an all-encompassing interpretation. In this respect, Wuthering Heights shares a feature of Hamlet.

**1. According to the passage, which of the following is a true statement about the first and second parts of *Wuthering Heights*?**

A The second part has received more attention from critics.
B The second part has little relation to the first part.
C The second part annuls the force of the first part.
D The second part provides less substantiation for a romantic reading.
E The second part is better because it is more realistic.

Many critics of Emily Brontë's novel Wuthering Heights see its second part as a counterpoint that comments on, if it does not reverse, the first part, where a romantic reading receives more confirmation. Seeing the two parts as a whole is encouraged by the novel's sophisticated structure, revealed in its complex use of narrators and time shifts. Granted that the presence of these elements need not argue for an authorial awareness of novelistic construction comparable to that of Henry James, their presence does encourage attempts to unify the novel's heterogeneous parts. However, any interpretation that seeks to unify all of the novel's diverse elements is bound to be somewhat unconvincing. This is not because such an interpretation necessarily stiffens into a thesis (although rigidity in any interpretation of this or of any novel is always a danger), but because Wuthering Heights has recalcitrant elements of undeniable power that, ultimately, resist inclusion in an all-encompassing interpretation. In this respect, Wuthering Heights shares a feature of Hamlet.

**Your answer is correct! Please click here go to the next question.**

**1. According to the passage, which of the following is a true statement about the first and second parts of *Wuthering Heights*?**

A The second part has received more attention from critics.
B The second part has little relation to the first part.
C The second part annuls the force of the first part.
D **The second part provides less substantiation for a romantic reading.**
E The second part is better because it is more realistic.

Many critics of Emily Brontë's novel Wuthering Heights see its second part as a counterpoint that comments on, if it does not reverse, the first part, where a romantic reading receives more confirmation. Seeing the two parts as a whole is encouraged by the novel's sophisticated structure, revealed in its complex use of narrators and time shifts. Granted that the presence of these elements need not argue for an authorial awareness of novelistic construction comparable to that of Henry James, their presence does encourage attempts to unify the novel's heterogeneous parts. However, any interpretation that seeks to unify all of the novel's diverse elements is bound to be somewhat unconvincing. This is not because such an interpretation necessarily stiffens into a thesis (although rigidity in any interpretation of this or of any novel is always a danger), but because Wuthering Heights has recalcitrant elements of undeniable power that, ultimately, resist inclusion in an all-encompassing interpretation. In this respect, Wuthering Heights shares a feature of Hamlet.

**Your answer is incorrect. Can you please read the question again?**

**1. According to the passage, which of the following is a true statement about the first and second parts of *Wuthering Heights*?**

A The second part has received more attention from critics.
B The second part has little relation to the first part.
C The second part annuls the force of the first part.
D The second part provides less substantiation for a romantic reading.
E The second part is better because it is more realistic.

Many critics of Emily Brontë's novel Wuthering Heights see its second part as a counterpoint that comments on, if it does not reverse, the first part, where a romantic reading receives more confirmation. Seeing the two parts as a whole is encouraged by the novel's sophisticated structure, revealed in its complex use of narrators and time shifts. Granted that the presence of these elements need not argue for an authorial awareness of novelistic construction comparable to that of Henry James, their presence does encourage attempts to unify the novel's heterogeneous parts. However, any interpretation that seeks to unify all of the novel's diverse elements is bound to be somewhat unconvincing. This is not because such an interpretation necessarily stiffens into a thesis (although rigidity in any interpretation of this or of any novel is always a danger), but because Wuthering Heights has recalcitrant elements of undeniable power that, ultimately, resist inclusion in an all-encompassing interpretation. In this respect, Wuthering Heights shares a feature of Hamlet.

**Your answer is incorrect. The question wants you to compare the first and second part of the novel.**

**1. According to the passage, which of the following is a true statement about the first and second parts of *Wuthering Heights*?**

A The second part has received more attention from critics.
B The second part has little relation to the first part.
C The second part annuls the force of the first part.
D The second part provides less substantiation for a romantic reading.
E The second part is better because it is more realistic.

Many critics of Emily Brontë's novel Wuthering Heights see its second part as a counterpoint that comments on, if it does not reverse, the first part, where a romantic reading receives more confirmation. Seeing the two parts as a whole is encouraged by the novel's sophisticated structure, revealed in its complex use of narrators and time shifts. Granted that the presence of these elements need not argue for an authorial awareness of novelistic construction comparable to that of Henry James, their presence does encourage attempts to unify the novel's heterogeneous parts. However, any interpretation that seeks to unify all of the novel's diverse elements is bound to be somewhat unconvincing. This is not because such an interpretation necessarily stiffens into a thesis (although rigidity in any interpretation of this or of any novel is always a danger), but because Wuthering Heights has recalcitrant elements of undeniable power that, ultimately, resist inclusion in an all-encompassing interpretation. In this respect, Wuthering Heights shares a feature of Hamlet.

**Your answer is incorrect. Can you please pay attention to the underlined sentences in the text?**

**1. According to the passage, which of the following is a true statement about the first and second parts of *Wuthering Heights*?**

A The second part has received more attention from critics.
B The second part has little relation to the first part.
C The second part annuls the force of the first part.
D The second part provides less substantiation for a romantic reading.
E The second part is better because it is more realistic.

Many critics of Emily Brontë's novel Wuthering Heights see its second part as a counterpoint that comments on, if it does not reverse, the first part, where a romantic reading receives more confirmation. Seeing the two parts as a whole is encouraged by the novel's sophisticated structure, revealed in its complex use of narrators and time shifts. Granted that the presence of these elements need not argue for an authorial awareness of novelistic construction comparable to that of Henry James, their presence does encourage attempts to unify the novel's heterogeneous parts. However, any interpretation that seeks to unify all of the novel's diverse elements is bound to be somewhat unconvincing. This is not because such an interpretation necessarily stiffens into a thesis (although rigidity in any interpretation of this or of any novel is always a danger), but because Wuthering Heights has recalcitrant elements of undeniable power that, ultimately, resist inclusion in an all-encompassing interpretation. In this respect, Wuthering Heights shares a feature of Hamlet.

**Your answer is incorrect. Can you please examine the underlined part in the text?**

**1. According to the passage, which of the following is a true statement about the first and second parts of *Wuthering Heights*?**

A The second part has received more attention from critics.
B The second part has little relation to the first part.
C The second part annuls the force of the first part.
D The second part provides less substantiation for a romantic reading.
E The second part is better because it is more realistic.

Many critics of Emily Brontë's novel Wuthering Heights see its second part as a counterpoint that comments on, if it does not reverse, the first part, where a romantic reading receives more confirmation. Seeing the two parts as a whole is encouraged by the novel's sophisticated structure, revealed in its complex use of narrators and time shifts. Granted that the presence of these elements need not argue for an authorial awareness of novelistic construction comparable to that of Henry James, their presence does encourage attempts to unify the novel's heterogeneous parts. However, any interpretation that seeks to unify all of the novel's diverse elements is bound to be somewhat unconvincing. This is not because such an interpretation necessarily stiffens into a thesis (although rigidity in any interpretation of this or of any novel is always a danger), but because Wuthering Heights has recalcitrant elements of undeniable power that, ultimately, resist inclusion in an all-encompassing interpretation. In this respect, Wuthering Heights shares a feature of Hamlet.

**The correct answer is D. Please click here go to the next question.**

**1. According to the passage, which of the following is a true statement about the first and second parts of *Wuthering Heights*?**

A The second part has received more attention from critics.
B The second part has little relation to the first part.
C The second part annuls the force of the first part.
D **The second part provides less substantiation for a romantic reading.**
E The second part is better because it is more realistic.

**CURRICULUM VITAE**

**Work Experience:**

| | |
|---|---|
| **2021-Present** | Ören Şehit Nusret Akar Secondary School (English Language Teacher) |
| **2019-2020** | Ören Şehit Nusret Akar Primary School (English Language Teacher) |

**Education:**

| | |
|---|---|
| **Master's Degree:** | Kocaeli University English Language Teaching (2018-2021) |
| **Bachelor's Degree:** | Hacettepe University English Language Teaching (2014-2018) |
| **High School:** | Balıkesir İstanbulluoğlu Anatolian Teacher Training High School (2010-2014) |

**Certifications:**

| | |
|---|---|
| **2019** | International Marmara Social Sciences Congress (Paper Presentation) |
| **2018** | Hacettepe University English Language Teaching Undergraduate Student Conference (Conference Volunteer) |
| **2018** | METU 12th Linguistics Conference |
| **2018** | Gazi University 2nd GELTUS English Language Teaching Undergraduate Student Conference |