

INFLUENCE OF AN AUDIENCE-ORIENTED TASK, EXECUTIVE FUNCTION,
AND TRAIT PERSPECTIVE TAKING ON EFL WRITING

TURGUT COŞKUN

BOĞAZİÇİ UNIVERSITY

2023

INFLUENCE OF AN AUDIENCE-ORIENTED TASK, EXECUTIVE FUNCTION,
AND TRAIT PERSPECTIVE TAKING ON EFL WRITING

Thesis submitted to the
Institute for Graduate Studies in Social Sciences
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Foreign Language Education

by
Turgut Coşkun

Boğaziçi University

2023

DECLARATION OF ORIGINALITY

I, Turgut Coşkun, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature

Date

ABSTRACT

Influence of an Audience-Oriented Task, Executive Function, and Trait Perspective Taking on EFL Writing

This study adopted a cognitive approach and investigated the effect of an audience-oriented task on the passages composed by novice EFL writers by considering their executive function and trait perspective-taking capabilities. With this purpose, the writing performance of two groups assigned to either an experimental or a control condition was compared. The audience-oriented experimental group was exposed to a writing task that emphasized the audience. Most importantly, they were asked to read three short messages shared by their audience, then take their perspective and write down whatever came to their minds. The non-audience-oriented control group was also asked to read the messages but was not encouraged to take the audience's perspective. They were just asked to write down whatever came to their minds. Following that, both groups wrote the main passages and completed four executive function tasks and a trait perspective-taking scale. The results revealed that exposing upper-intermediate novice EFL writers to an audience-oriented task increased the overall writing quality of the participants with high executive function or trait perspective-taking capabilities. Regarding the level of persuasiveness, the task's effect may depend on the availability of sufficient executive function resources. These findings confirm the importance of a task for directing and managing cognitive resources and show the executive function's central role and trait perspective-taking's importance in writing. The tasks that encourage employing these individual resources may enrich the instructor toolboxes.

ÖZET

Okuyucu Odaklı Yazma Görevi, Yürütücü İşlev ve

Bakış Açısı Alma Özelliklerinin İngilizce Yazma Performansı Üzerine Etkisi

Bu çalışmada, bilişsel yaklaşım temelinde oluşturulan okuyucu odaklı bir görevin, yürütücü işlev ve bakış açısı alma özellikleri de dikkate alınarak, İngilizce öğrenen üst-orta seviye deneyimsiz yazarların oluşturduğu metinler üzerindeki etkisi, deney veya kontrol gruplarına rastgele atanan katılımcıların yazma performansı karşılaştırılarak araştırıldı. Deney grubunda bulunan katılımcılar okuyucunun daha fazla dikkate alındığı yazma görevine maruz bırakıldı. En önemlisi, bu gruptakiler okuyucuları tarafından oluşturulan üç farklı mesajın her birini okuduktan sonra bu okuyucuların bakış açısını dikkate aldı ve akıllarına ne gelirse üç dakika boyunca yazdı. Buna karşın kontrol grubunda bulunanlar mesajları okuduktan sonra okuyucuların bakış açılarını dikkate almadan sadece akıllarına gelenleri yazdılar. Sonrasında her iki grup da ana metinleri yazıp dört yürütücü işlev testini ve çevrimiçi verilen bakış açısı ölçeğini tamamladılar. Sonuçlar verilen okuyucu odaklı yazma görevinin yürütücü işlev ve bakış açısı alma ile etkileşime girdiği ve yüksek seviyede yürütme işlevi veya bakış açısı alma becerisi olan katılımcıların performansını arttırdığı gösterdi. Metinlerin ikna edicilik düzeyi dikkate alındığında ise yazma görevi etkisinin yüksek seviyede yürütme işlevi kapasitesinin varlığına bağlı olabileceği görüldü. Bulgular bilişsel kaynakların yönlendirilme ve idaresinde yazma görevinin önemini teyit etmiş, yürütücü işlevin merkezi konumunu ve bakış açısı alma özelliğinin önemini göstermiştir. Bu tarz bireysel kaynakları kullanmaya teşvik eden yazma görevleri eğitmenlerin kaynaklarını zenginleştirebilir.

CURRICULUM VITAE

NAME: Turgut Coşkun

DEGREES AWARDED

PhD in Foreign Language Education, 2023, Boğaziçi University

MA in Cognitive Psychology, 2014, Boğaziçi University

BA in English Literature, 2010, Doğuş University

AREAS OF SPECIAL INTEREST

Executive function, memory, attention, second language learning, vocabulary processing, sentence processing, writing, audience processing

PROFESSIONAL EXPERIENCE

English Instructor, Preparatory School, Kadir Has University, 2019 - present

English Instructor and Planning Coordinator, Preparatory School,
İstanbul Aydın University, 2017 - 2019

English Instructor, Preparatory School, İstanbul Aydın University, 2017

AWARDS AND HONORS

High Honors, Doğuş University, 2010

PUBLICATIONS

Master Thesis

Coskun, T. (2014) Görsel bilginin çalışan belleğe alınma ve işlenme sırası. Yüksek lisans tezi, danışman: Dr. Aysecan Boduroglu, Bilişsel Psikoloji, Boğaziçi üniversitesi, Türkiye. (Ulusal Tez Merkezi TEZ NO: 361951).

Conference Proceedings

Coşkun, T., Boduroglu, A. (August 2016). Temporal processing of visual information and its influence on visual working memory representation, ECVP 2016, Barcelona.

ACKNOWLEDGEMENTS

Thanks to Assoc. Prof. Nur Yiğitoğlu Aptoula not only for increasing my interest in writing studies via her valuable lectures but also for adapting a humanistic approach and creating a psychologically safe environment for my study. I always felt her guidance, encouragement, and support. I thank Assoc. Prof. Senem Yıldız for her intellectually entertaining lectures and comments. Besides many things, they helped me internalize the importance of the task in language instruction and contributed to my intellectual development. I want to thank Professor Çiler Hatipoğlu for her academic studies and detailed feedback, which helped me better understand Turkish EFL learners' writing processes. I thank Gülcan Erçetin for her mind-opening lectures, which raised my attention to automatic versus controlled processes and established a background for this study. I also thank Professor Ayşe Gürel for her enthusiastic psycholinguistic lectures, which established a connection between my cognitive psychology background and EFL studies; I always felt her sincere academic support. Again, I thank Professor Ümit Deniz Turan for her guidance and highly intellectual and enjoyable lectures, which brought my academic curiosity back to life. Unfortunately, it is not possible to mention all the names, but I thank all my teachers, from my dear primary school teacher Eda Durdu to the Ph.D. level Professors Nazik Deniz Dinçtopal and Sibel Tatar.

Again, I thank my dear friends and colleagues who contributed to this study as second coders and evaluators (Alena Dhahi, Gülsen Yılmaz, Gülşah Sobucalı, Mehtap Güven, Onur Keleş, and Oylum Savlak) and those who helped me transcribe a part of the data (Ahmet Ermeç, Aleyna Söylemez, Melodi Korkmaz, Muhammed Ermeç, and

Neşe Şenol). I also thank those Kadir Has University Preparatory School's sincere instructors and management for their kind support during the data collection.

Again, I thank all the researchers who contributed to this study via their studies and especially those who kindly spared their time to give their valuable feedback when I encountered some challenging situations (Cemre Çiçek Tümer, Engle's Lab members, Ken Hyland, and many others).

Lastly, I want to thank all my dear students, including those who voluntarily contributed to pilot studies. Beyond that, the students who participated in this study were always supportive during and after the data collection. I sincerely thank each of them. Moreover, most others were so kind and intellectually promising that I felt I must contribute by conducting studies that could improve their performance. I want to write all the names, but I do not want to risk missing any of them. Having those kinds of students was the real driving force behind this study. Thanks to them, I have almost never felt regretful for giving up and sacrificing things, material and nonmaterial.

DEDICATION

To Zeynep and Feyza.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION	1
1.1 The research background	1
1.2 The aim and possible contribution of the present study	4
1.3 The organization of the thesis	4
CHAPTER 2: COGNITIVE WRITING PROCESSES	6
2.1 A pioneering cognitive process model	6
2.2 Involvement of executive function and audience in the following models	9
CHAPTER 3: DIRECTING ATTENTION TO AUDIENCE	34
3.1 Language processing	34
3.2 Audience processing	37
3.3 Designing a task directing the EFL writer's attention to the audience	41
CHAPTER 4: PRESENT STUDY	48
4.1 The theoretical and empirical background	48
4.2 The aim and contribution	51
4.3 The research questions	52
CHAPTER 5: METHODOLOGY	54
5.1 Participants	54
5.2 Context	55
5.3 Materials	62
5.4 Procedure	76

5.5 Target writing and control measurements	82
5.6 Ethical concerns	103
CHAPTER 6: RESULTS	105
6.1 Data preparation and preliminary analyses	106
6.2 Main analyses	139
6.3 Additional analyses	157
CHAPTER 7: DISCUSSION	163
7.1 Conclusion	164
7.2 Further discussion and future directions	194
7.3 Pedagogical implications	203
7.4 Contribution	211
7.5 Limitations	211
APPENDIX A: ETHICAL COMMITTEE APPROVAL FORM	216
APPENDIX B: PERSONAL INFORMATION FORM	217
APPENDIX C: DIAGNOSTIC WRITING TASK	219
APPENDIX D: MAIN WRITING TASKS	220
APPENDIX E: THE IMAGES USED IN THE EXPERIMENT	234
APPENDIX F: RETROSPECTIVE QUESTIONNAIRE	237
APPENDIX G: PERSPECTIVE TAKING SCALE	242
APPENDIX H: WRITING ASSESSMENT (OVERALL WRITING QUALITY) SCALE	243
APPENDIX I: THE LEVEL OF PERSUASIVENESS RUBRIC.....	244

APPENDIX J: POST WRITING (BACKGROUND) CODEGUIDE	253
APPENDIX K: POST WRITING (REFLECTIVE) CODEGUIDE	268
APPENDIX L: CODING STRATEGY USE DURING EF TASKS	277
APPENDIX M: CODING/EVALUATING SIMPLE BRAINSTORMING AND PERSPECTIVE TAKING NOTES	285
APPENDIX N: CODING INTERACTIONAL MARKERS	287
APPENDIX O: CODING/SEPARATING T-UNITS	291
REFERENCES	293

LIST OF TABLES

Table 1. The Use of Working Memory Resources by Six Basic Processes	
Writing	19
Table 2. The Distribution of Participant Gender and Department Types across	
Groups	55
Table 3. The Distribution of Assessment Components	57
Table 4. A Sample Possibility Table for a B1 Level Student	57
Table 5. The Conditions to Complete the Preparatory School Successfully	58
Table 6. The Distribution of the Initial Levels across Groups	59
Table 7. The Material List	63
Table 8. Retrospective Writer Background Questions and the Interrater	
Reliability Values	97
Table 9. Retrospective Writing Reflection Questions and the Interrater	
Reliability Values	98
Table 10. Questions for the Strategy Use in EF Tasks and the Interrater	
Reliability Values	100
Table 11. The Median Z-Values of AC, WM, and EF	114
Table 12. The Distribution of those who Took and did not Take English and	
Turkish Writing Lessons Before the Preparatory School	116
Table 13. The Distribution of the Reported Difficulty in General English and	
Turkish Writing	117

Table 14. The Distribution of the Coded Sources of Reported Difficulties in English and Turkish Writing	118
Table 15. The Distribution of the Reported Thinking Language while Writing in English	119
Table 16. The Distribution of the Writing Activities Done during or after the Lessons	121
Table 17. The Distribution of the Reported Interest in Music, Theatre, and/Outdoor Sports Activities and Clubs across the Groups	124
Table 18. The Perceived Difficulty of the Prompt across the Groups	125
Table 19. The Perceived Difficulty of Main Writing across the Groups	126
Table 20. The Percentage of Difficulty in General Writing Compared to the Difficulty in Main Writing	127
Table 21. The Reported Sources of Difficulty in Main Writing	128
Table 22. The Review Pattern of the Groups	129
Table 23. The Predictions Made about the Evaluation of Main Writing	131
Table 24. The Distribution of the Reported Audience PT Across the Groups ...	134
Table 25. The Reported Influence of Audience PT on Writing across the Groups	134
Table 26. The Standardized Frequency of Engagement and Self-Mention Markers Across the Experimental and Control Groups	138
Table 27. Overall Writing Quality Performance of Higher EF Participants Predicted by Experimental Task Condition Controlling the Language and Trait PT Background	145

Table 28. Overall Writing Quality Performance of the Higher Trait PT Group Predicted by Experimental Task Condition Controlling the Language and EF Background	151
Table 29. The Mean Overall Writing Quality Values as a Function of EF, trait PT, and Experimental Conditions	155
Table 30. The Mean Level of Persuasiveness Values as a Function of EF, trait PT, and Experimental Conditions	157
Table 31. The Associations between Three Critical Variables of the Study: Diagnostic Overall Writing Quality, EF, and Trait PT	161

LIST OF APPENDIX TABLES

Table J1. The Categories and Samples for Taking English Writing Lessons	
Before	253
Table J2. The Categories and Samples for General English Writing	
Performance	255
Table J3. The Categories and Samples for Subcategories of Experiencing	
Difficulty in General English Writing	256
Table J4. The Categories and Samples for Taking Turkish Writing Lessons	
Before	258
Table J5. The Categories and Samples for General Turkish Writing	
Performance	259
Table J6. The Categories and Samples for Subcategories of Experiencing	
Difficulty in General Turkish Writing	260
Table J7. The Categories and Samples for Thinking in Turkish or in English	
while Writing in English	262
Table J8. The Categories and Samples for Writing Activities Done during or	
after the Lessons	263
Table J9. The Categories and Samples for Interest in Music, Theatre, and	
Outdoor Sports	264
Table J10. The Categories and Samples for Being Active in Music, Theatre,	
and Outdoor Sports	266
Table K1. The Categories and Definitions for Taking Audience Perspective	269

Table K2. The Categories and Definitions for the Influence of Taking Audience Perspective	271
Table K3. The Categories and Samples for the Perceived Difficulty of the Prompt	272
Table K4. The Categories and Samples for the Perceived Difficulty in the Main Writing	273
Table K5. The Categories and Samples for the Sources/Subcategories of the Perceived Difficulty in the Main Writing	274
Table K6. The Categories and Samples for Reviewing after Writing	275
Table K7. The Categories and Samples for the Predictions Made about the Evaluation of Main Writing	276
Table L1. The Categories and Samples for the Reported Strategy Use in Flanker Task	278
Table L2. The Categories and Samples for the Reported Strategy Use in Antisaccade Task	280
Table L3. The Categories and Samples for the Reported Strategy Use in Symmetry Span Task	282
Table L4. The Categories and Samples for the Reported Strategy Use in Foster Operational Span Task	284
Table N1. Self-Mention and EM Categories	288

LIST OF FIGURES

Fig. 1 The writing model introduced by Flower and Hayes (1981)	7
Fig. 2 The revised WM model of Baddeley (2000)	11
Fig. 3 The demonstration of the activation of long-term memory and the involvement of executive attention	13
Fig. 4 The revised model of Hayes	15
Fig. 5 The revised cognitive writing processes model	21
Fig. 6 Direct and indirect effects model of writing	27
Fig. 7 Expanded direct and indirect effects model of writing	30
Fig. 8 A sample anti-saccade task trial	73
Fig. 9 A sample symmetry span task trial	74
Fig. 10 A sample operational span task trial	75
Fig. 11 A summary of the data collection procedure	81
Fig. 12 The overall writing quality performance of relatively higher and lower EF groups in the experimental and control conditions	142
Fig. 13 The level of persuasiveness performance of relatively higher and lower EF groups in the experimental and control conditions	146
Fig. 14 The overall writing quality performance of relatively higher and lower PT groups in the experimental and control task conditions	149
Fig. 15 The level of persuasiveness performance of relatively higher and lower PT groups in the experimental and control task conditions	153

Fig. 16 The overall writing quality pattern revealed by relatively low PT - low EF group across their within-subject diagnostic versus main writing and between subject experimental versus control conditions	160

LIST OF APPENDIX FIGURES

Fig. E1 The outdoor sports club visual used for audience-oriented experimental task	234
Fig. E2 The theatre club visual used for audience-oriented experimental task ...	234
Fig. E3 The music club visual used for audience-oriented experimental task	235
Fig. E4 The outdoor sports club visual used for non-audience-oriented control task	235
Fig. E5 The theatre club visual used for non-audience-oriented control task	236
Fig. E6 The music club visual used for non-audience-oriented control task	236

ABBREVIATIONS

AC	Attention control (or executive attention)
DIEW	Direct and Indirect Effects Model of Writing
EF	Executive function
EFL	English as a foreign language
EM	Engagement (metadiscourse) marker
LTM	Long-term memory
L1	First language
L2	Second language
PT	Perspective taking
SM	Self-mention (metadiscourse) marker
WM	Working memory

INTENTIONALLY LEFT BLANK

CHAPTER 1

INTRODUCTION

Writing is a highly complex task requiring many language and cognitive processes and their coordination to address an audience who is not immediately available. I will get insight into some of these processes in this study. I will adopt a cognitive theoretical approach, consider executive function and trait perspective taking, and try to understand the influence of a writing task that directs the attention of novice Turkish B2 (upper-intermediate) level EFL writers to their readers.

This section will start with the introduction of the research background, continue with the aim and possible contribution, and finish with the organization of the thesis.

1.1 The Research background

Since the 1980s, many models have been introduced to explain cognitive processes involved in writing (e.g., Hayes & Flower, 1980a; Hayes, 1996; Kellogg, 1996; Kim & Park, 2019). One of the most important characteristics of these cognitive models was recognizing the possible contribution of the limited executive function (EF) resources, which are suggested to be responsible for the regulation and control of cognitive processes (Miyake et al., 2000) in writing (e.g., Becker, 2006; Hayes, 1996; Kellogg, 1996; Kim & Park, 2019). The importance of audience awareness (e.g., considering the readers' needs) was also emphasized in many of them.

For example, Direct and Indirect Effects Model of Writing (DIEW) (Kim & Park, 2019) suggests that writing activity needs to generate ideas, translate them into

oral language, and transcribe them into a written text. As this happens, many components in hierarchical structural relations are involved. At the lowest level, the limited EF resources, which include working memory (WM) and attention control (AC), establish a necessary domain-general cognitive foundation for the basic requirements, such as vocabulary, grammar, and transcribing. However, as the model suggests, writing activity consists of more than these basic language processes. It also employs higher-level processes such as perspective-taking (PT), which is directly associated with audience awareness and defined as “one’s knowledge of and inferences about others’ mental and emotional states in writing” (p. 1323), and which also requires limited EF resources. Hence, both the primary language processes and the higher-level processes, such as considering the audience while writing, may recruit this limited cognitive resource, and this may create competition.

The effect of EF demand and the possible competition between lower and higher-level writing components may be more visible, particularly if an EFL learner's language proficiency is not high enough. This may be because these second language (L2) writers need, for example, to transform nonverbal messages, such as propositional content/ideas, into appropriate verbal-linguistic forms. While doing this, compared to higher (advanced) level EFL learners, relatively lower (e.g., upper-intermediate) level EFL learners can have more difficulties, for instance, in connecting parts of a text appropriately (Schoonen, Van Gelderen, Stoel, Hulstijn, & De Glopper, 2011) and these low-level processes can dominate the limited EF resources which are required to orchestrate writing by allocating attention most efficiently (Kormos, 2012; Weigle,

2005). As a result, novice EFL writers may consider the higher-level processes less, and their writing performance may be decreased.

It may surely depend on the availability of EF resources (McCutchen, 1996), but one way to encourage a low-proficiency novice EFL writer's cognitive system to integrate higher-level audience processes and use EF resources more efficiently may be via adapting a writing task that directs their attention to reader perspectives. The studies indicated that the existence of an audience can be effectively emphasized in a single task, and doing this can result in higher audience awareness and better passage adaptation (e.g., Block & Strachan, 2019; Cho & Choi, 2018; Cohen & Riel, 1989). Hence, although it can be demanding (Kim, 2020; Kim & Graham, 2021; Ryskin, Benjamin, Tullis, & Brown-Schmidt, 2015; Wardlow, 2013), a well-designed task can successfully trigger audience-oriented processes and keep them maintained while writing.

Importantly, besides having higher EF capability, having a higher trait PT ability (a natural/general tendency to take other people's perspectives) may contribute to the success of this task. This is because the studies indicated that those people with higher trait PT could, for example, judge other people's unique level of traits better (Colman, Letzring, & Biesanz, 2017), and they can perform better when they are asked to take the perspective of another person in different contexts (Crocetti et al., 2016). Moreover, having such a natural tendency may make the state (situational) reader PT performance less dependent on EF resources and compensate for the lack of sufficient EF resources.

1.2 The aim and possible contribution of the present study

This study aims to evaluate the validity of these suggestions within a Turkish EFL learning context. I will create an experimental audience-oriented writing task that includes some instruments (e.g., PT activity) which can direct the writers' attention to the audience. A non-audience-oriented task that does not employ these instruments will function as a control. The question is whether the experimental task can increase the writing performance of upper-intermediate level novice EFL writers by interacting with EF capability (low versus high) or trait PT (low versus high)¹. Lastly, I will consider whether having higher trait PT may compensate for the lack of sufficient EF resources.

The previous studies indicated the importance of EF (e.g., WM) (Vasylets & Marin, 2021) and PT (Cho, Kim, & Olson, 2021), which is associated with audience awareness (Kim & Park, 2019). However, to my knowledge, no previous studies created a writing task that directed writers' attention to the audience and investigated its effect on writing performance by considering the contribution of EF capability or trait PT tendencies. This study may fill this gap by investigating whether an audience-oriented writing task interacts with EF or trait PT and increases the writing performance of upper-intermediate level novice EFL writers.

1.3 The organization of the thesis

The organization of the thesis is as the following. Chapter 2 focuses on cognitive processes involved in writing and reviews cognitive writing models. Chapter 3 introduces some background for the task that aims to direct EFL writers' attention to the

¹ Participants who tend to take other people's perspectives relatively more or less.

audience. Chapter 4 focuses on the rationale behind the present study and introduces the research questions. Chapter 5 focuses on the methodology (including the research context, materials, and measurements) of the study. Chapter 6 analyzes the data and reports the results. It starts with data preparation and preliminary analyses and then focuses on the main analyses, which investigate the research questions. The chapter finishes with some additional analyses. Chapter 7 shares the conclusions derived from the results and discusses them by suggesting future directions. It also shares some pedagogical implications for instructors, considers the study's possible contribution, and recognizes its possible limitations.

CHAPTER 2

COGNITIVE WRITING PROCESSES

As mentioned above, the theoretical basis of this study will be founded upon cognitive writing processes literature. With this purpose in mind, I will start this chapter by reviewing a pioneering cognitive model, and then after introducing the EF concept, I will focus on more recent models which directly emphasize the importance of EF and audience.

2.1 A pioneering cognitive process model

One of the earliest and most efficient models which tried to explain the role of cognitive processes in writing was introduced by Hayes and Flower (1980a). The basic underlying suggestion was that “writers are constantly, instant by instant, orchestrating a battery of cognitive processes as they integrate planning, remembering, writing, and rereading” (Flower & Hayes, 1981, p. 387).

Although many further subprocesses are embedded under them, the model divides the writer’s world into three main components: task environment, long-term memory (LTM) of the writer, and writing processes (Fig. 1) (Flower & Hayes, 1981; Hayes & Flower, 1980a).

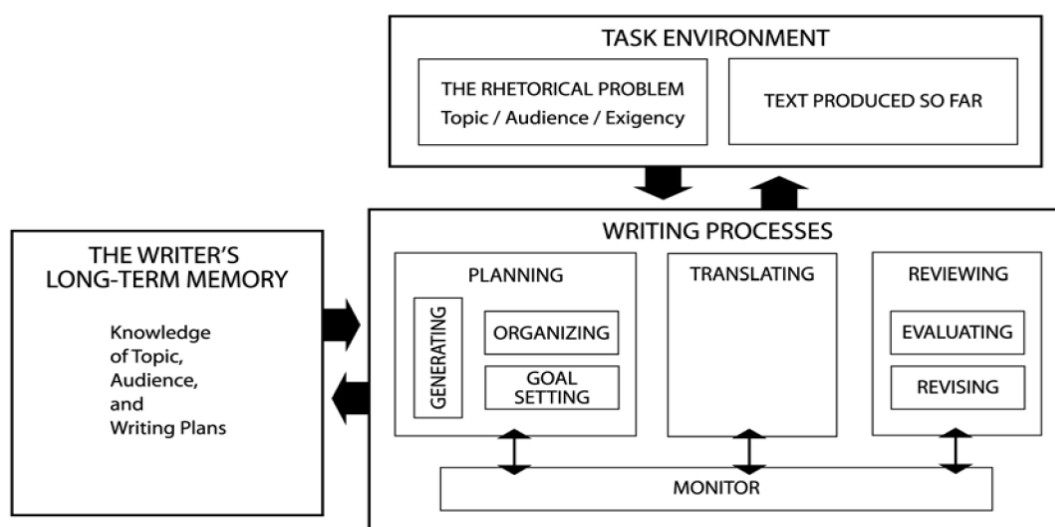


Fig. 1 The writing model introduced by Flower and Hayes (1981)

Taken from Flower and Hayes (1981, p. 370).

The first component, the task environment, includes all the things outside the cognition of a writer. It involves two subcomponents. One is the rhetorical problem (or writing task) which consists of the topic of a task and audience (who is going to read and evaluate the written text), as well as exigency (the writer's aim for writing). The other is the growing text itself. Like the other task constraints, each word written until that moment, particularly if the text is coherent, limits the writer's choices for what to say next. In this way, the growing text increases the time and attention demand for a writer.

The second component, LTM², is a permanent storage system that stores topic background, audience, writing plans, etc. (Hayes & Flower, 1980a). This memory

² Flower and Hayes (1981) define LTM as a stable entity with its own internal organization and differentiate it from WM, defined as conscious attention or active processing capacity. Although they do not consider WM in this early model, they will recognize its importance and add it to their later models (see below).

system is involved in the solutions of writing problems, such as using a cue to retrieve useful knowledge and adapting that retrieved knowledge to solve a rhetorical problem or task.

The third component of the model is the writing processes which are under the control of the monitor³ (Hayes & Flower, 1980a). These processes are planning, translating, and reviewing. Planning includes three further subprocesses: generating, organizing, and goal-setting (Hayes & Flower, 1980a). The planning process considers the task environment and the information retrieved from LTM for setting goals and organizing a final writing plan to guide the text. The planning here is an internal representation that can be visual or perceptual rather than a writer's final prose representation; thus, it is an abstract concept (Flower & Hayes, 1981). The first subprocess of planning, generating ideas, includes the retrieval of relevant information from LTM. Sometimes, this retrieval process may be pretty smooth, but it may be less successful at other times. For example, the writer may retrieve fragmentary or contradictory ideas. In this case, another subprocess, organizing, can help the writer bring pieces together and create a meaningful whole (Flower & Hayes, 1981). Moreover, thanks to the organizing process, ideas may be grouped (subordinate-superordinate relations, etc.), and new concepts may be formed. This organizing is generally guided by goal-setting, which is another important aspect of planning (Flower & Hayes, 1981). The second writing process is translating⁴, which functions for

³ The monitor determines the process and progress during the composition, for example, how much time will be spent to generate ideas before starting to write, etc.

⁴ They purposefully use the term translating to differentiate it from simply writing or transcribing because, according to them, this process can include the translation of symbols, kinetics, visuals, etc., into a written language.

translating meaning into written language. As Flower and Hayes (1981) reported, the translation process requires dealing with language formality, syntactic and lexical needs, forming letters, etc. If a writer is not experienced enough, the translating process may be an extra burden to the limited WM capacity. It can interfere with global processes, such as planning, or with local processes, such as sentence formation, or both. The last writing process is conscious reviewing, which includes evaluating and revising subprocesses. It can function during the revision/evaluation of both written and unwritten statements and can lead to new cycles of planning, etc. (Flower & Hayes, 1981).

These distinct cognitive processes, which are embedded with each other in a hierarchical organization, are orchestrated by writers while composing (Flower & Hayes, 1981). They do not progress linearly. For example, successfully reaching the audience of the text needs the organization of ideas at all levels, and although some previously established plans and goals may be retrieved from LTM, most of these plans and goals are generated, developed, and revised throughout composing. In other words, for example, goals lead to idea generation, but, in turn, these generated ideas can lead to more complex goals. Again, a writer can review, generate, and reorganize in the middle of composing.

2.2 Involvement of executive function and audience in the following models

Hayes and Flower's (1980a) seminal model, which identified the major cognitive processes in text production, created a backbone for many following models (Limpo & Olive, 2021). Important for this study's context, one of the most important

characteristics of these cognitive models is explicitly recognizing WM⁵ and executive attention⁶ as central processors in writing (e.g., Becker, 2006; Hayes, 1996; Kellogg, 1996; Kim & Park, 2019) and the other one is recognizing the importance of considering the audience. Before reviewing these models in detail, I will shortly give details about WM and executive attention concepts which are associated with EF.

The relatively earlier cognitive writing models adopted the operationalization of WM by Baddeley (1986), who defined it as a limited resource used to store information and carry out online cognitive processes and suggested that it includes three essential components. One is the phonological loop, a kind of inner voice that stores phonological information. The other is the visuospatial sketchpad, which deals with storing visual and spatial information. The third component is the central executive, which conducts cognitive tasks such as reasoning and deals with tasks such as memory retrieval or tasks that are not fully automated. In other words, if a process is not sufficiently automatized, it will use WM resources, and the more consciousness it involves, the more resources will be required (Baddeley, 1986).

Later, Baddeley (2000) revised this model by elaborating the interaction between three slave systems and a controlling central executive (Fig. 2). The comparatively passive slave systems are the phonological loop, visuospatial sketchpad, and the newly added episodic buffer. Beyond these, the central executive controls the whole system.

⁵ Some of these models use the term short-term memory rather than WM. Because the differences between these concepts are beyond the scope of this study and because I want to be consistent, I will use the term WM even when the original models (particularly those early ones) refer to the very same concept as short-term memory.

⁶ It can also be called AC; here, I used the term executive attention to be consistent with these papers, but they can be used interchangeably.

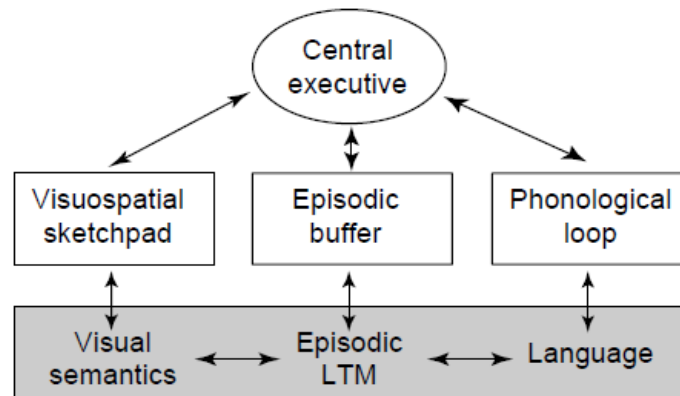


Fig. 2 The revised WM model of Baddeley (2000)

Taken from Baddeley (2000, p. 421).

The central executive (central/executive attention), which controls the slave systems, is the most active part of the system in this model. Episodic buffer interfaces between the visuospatial sketchpad, the phonological loop, and LTM while processing multi-dimensional codes. Thus, it is a place where different information types are integrated/bound to create coherence in information flow. Additionally, this model suggests that the episodic buffer is essential to transfer the learned episodes into LTM, which is defined as a place where permanent information is kept as, for example, visual semantics or language. Besides that, the central executive also controls the phonological loop, whose primary function is maintaining speech-based input by processing phonological and verbal information. This does not have to be directly spoken input. For example, an input may be presented visually, and this loop may process sub-vocal articulations of it. And the last slave system controlled by the central executive is the visuospatial sketchpad, whose primary function is to deal with visual and spatial information. The distinction between these components, for example, between

visuospatial sketchpad and phonological loop, is supported by some research. For instance, it was revealed that a reading span WM task is associated with verbal rather than visuospatial abilities. In contrast, a spatial span WM task is associated with visuospatial rather than verbal abilities (e.g., Shah & Miyake, 1996).

Baddeley's approach, which suggested the existence of different channels, for example, visual versus verbal (Baddeley & Hitch, 1974), were very influential. Still, some recent approaches do not adopt this kind of differentiation (Cowan, 2011). They mostly think of WM as a unitary concept that can be highly associated with executive attention (e.g., Engle, 2002; Engle, Tuholski, Laughlin, & Conway, 1999). In fact, both this approach and Baddeley's (2000) approach include a central executive component that directs attention and controls information processing, and both of them recognize the system's limited capacity. However, the latter approach suggests that the nature of short-term memory (storage components such as phonological loop) is different from WM and suggests that WM capacity can represent the limited ability to control attention which is particularly important to deal with the tasks which include conflicting/interfering distractors (Engle, 2002). Thus, WM capacity is not only about the number of items to be stored but also about the ability to control attention to maintain information in an active/retrievable state while suppressing/avoiding distractors.

Some other researchers who adopt this unitary approach consider WM as the focused part of LTM (Cowan, 2011). The suggestion is that some LTM representations can be activated. Within these activated representations, some portions (1-4 items) can be focused and processed by the attention mechanism, which can limit active

information processing: WM. Thus, in this approach, WM is considered a dynamic system that can transiently activate LTM representations (Fig. 3).

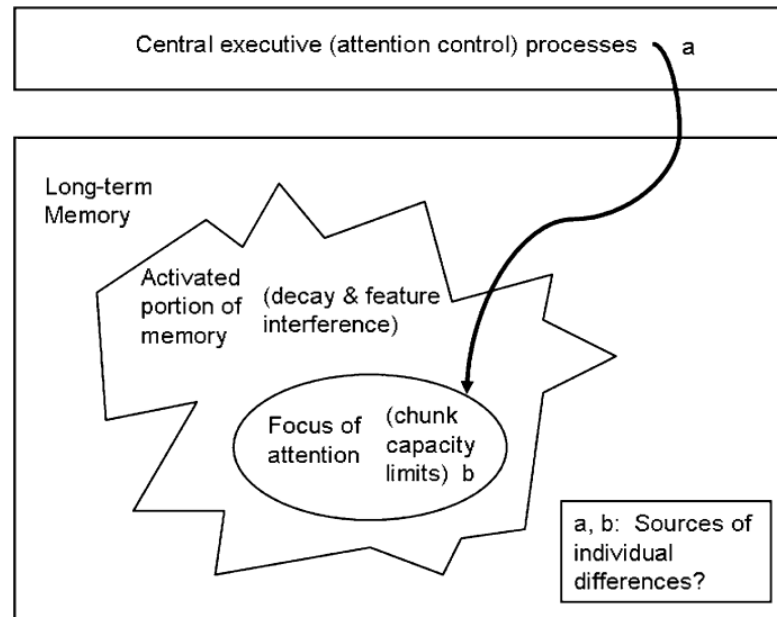


Fig. 3 The demonstration of the activation of LTM and the involvement of executive attention

Taken from Cowan (2008, p. 326).

Beyond these, an important concept related to WM and AC is EF. Some researchers define it as “high-level cognitive processes that facilitate new ways of behaving, and that optimize one’s approach to unfamiliar (non-routine) circumstances” (Gilbert & Burgess, 2008, p. 110) and suggest that it is “an umbrella term, referring to those high-level processes that control and organize other mental processes” (Gilbert & Burgess, 2008, p. 112). This account suggests EF includes core WM and attention capabilities (AC) such as inhibition (e.g., selective attention for controlling inference) and updating/monitoring WM representations (Diamond, 2013; Miyake et al., 2000). However, some other researchers, which I mentioned above (e.g., Engle, 2002), do not

differentiate these cognitive processes and suggest that executive attention (e.g., inhibitory mechanisms) and WM use the same system resources and depend on each other. For example, WM capacity determines how much an individual will be influenced by proactive interference (Engle & Kane, 2004). Hence, for them EF refers to both limited WM and executive attention (AC) resources.

In this study's context, I will adopt the definition of WM as mental processes responsible for online and temporary information processing and manipulation (Juffs & Harrington, 2011) and the unitary rather than the component-based approach. The term EF will be used to cover both executive attention (AC) and WM, which can be highly associated with each other.

Now that I have defined the basic concepts associated with EF, I will continue reporting the following cognitive writing approaches that explicitly embed EF resources (WM and AC) into the models.

In 1996, John Hayes considered social and motivational factors and elaborated cognitive processes in his revised individual-environment model. The revised model has two key components: the task environment and the individual (Fig. 4). Important to our present concern, a critical component of the task environment is the audience (e.g., the target reader of the passage), and a critical component of the individual is WM.

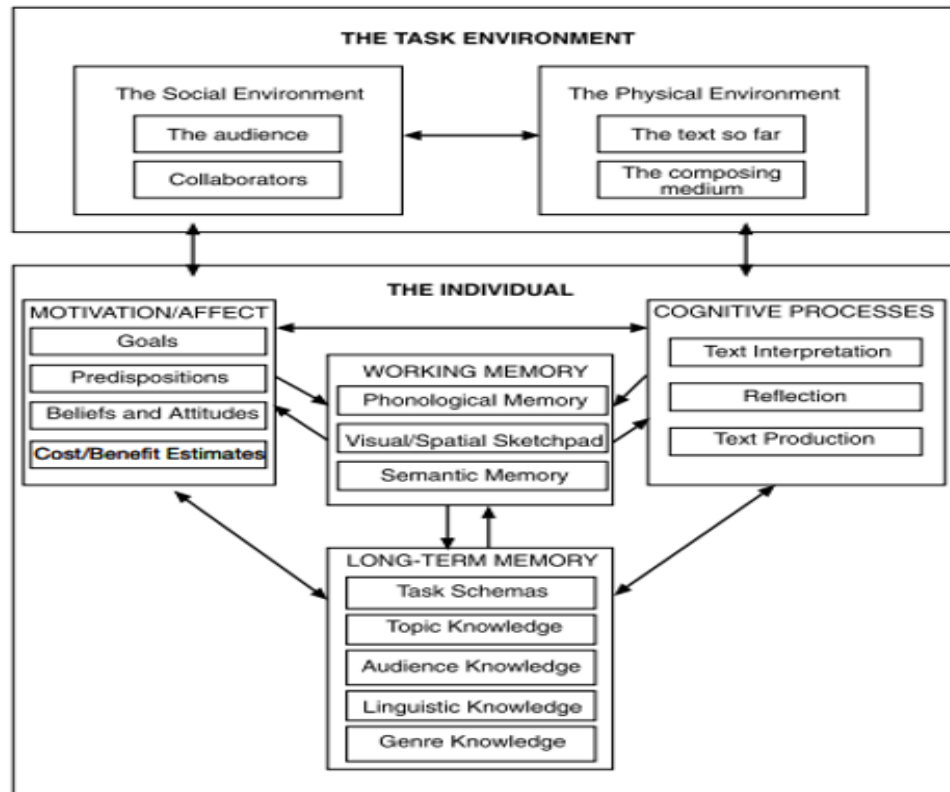


Fig. 4 The revised model of Hayes

Taken from Hayes (1996, p. 4).

To give more detail, the task environment is further divided into social and physical subcomponents. The social environment includes the audience and collaborators (e.g., the other passages which can be read). Hayes (1996) notes that writing is a social activity not only because it has a communicative purpose but also because it happens in a social setting that is shaped by a writer's previous interactions within a culture, etc. For example, a writer can compose differently for a familiar compared to an unfamiliar audience. Again, the physical environment, such as the text produced until that moment (e.g., the length of it and its level of coherence) and the writing medium (e.g., a word processor, an e-mail, etc.), can make a difference (Hayes, 1996).

The individual component comprises LTM, WM, motivation/affect, and cognitive processes (Hayes, 1996). LTM was kept with some updates/elaborations and embedded within the individual component. The previously constructed LTMs which can be retrieved during writing include task schemas (the processes to complete the task, etc.), topic knowledge, linguistic knowledge, genre knowledge (narrative, argumentative, etc.), and audience knowledge. All these memories have some functions. For example, writers can use audience knowledge to take a reader's perspective (a close friend versus a stranger, etc.) and adapt the passage accordingly. Apart from this elaboration, some important updates were reported in the individual components of the model. One of the most important updates under this component is exchanging the term monitor with the term WM. In the original model (Hayes & Flower, 1980a), the term monitor was used to specify a mechanism that controls basic writing processes (planning, translating, and reviewing). However, in the updated model, Hayes (1996) recognized WM's centrality in a writing activity⁷ and adopted Baddeley's WM model (1986). Another important update was the inclusion of motivation, which could be affected by the writer's goals, beliefs, attitudes, predispositions, and cost/benefit estimations. And lastly, cognitive processes were reorganized such that text interpretation replaced revision, planning was considered as only one of the reflection processes, and translation was considered a part of the text production process. Text interpretation is a function that creates internal representations from linguistic (listening and reading) and visual inputs (scanned graphics). Cognitive processes that work on internal representations to solve a problem, make a decision and inference, etc., are

⁷ Hayes (1996) noted that he considered planning and decision-making under reflection processes but semantic store under WM because the latter can help describe text generation.

called reflection. Hayes reports that the reflection process is important for readers and writers. Writers also make inferences to consider their readers' needs (the level of knowledge, etc.) and respond accordingly. Lastly, text production is a function that considers the context of a task and produces a passage (written or verbal) or a graphic output.

Hayes (1996) emphasized that these cognitive and motivational processes use WM and LTM resources that are not specific to writing; these resources are shared with other activities such as reading novels (text interpretation), solving puzzles (reflection), or drawing (production). Another thing that Hayes (1996) emphasized is the centrality of reading in this model: a) reading the written text so far to evaluate it, b) reading a source text to create a background for the topic and to respond to it⁸, and c) reading a task to understand its requirements. Lastly, he emphasized that all of the components and subcomponents in the model were essential. Writing depends on an appropriate combination of all these cognitive, affective, social, and physical conditions. This is because writing is a communicative and intellectual activity that requires a social context, a medium, motivation, and cognitive resources.

Another important writing model which integrated Hayes and Flower's (1980a) writing model and Baddeley's (1986) WM was introduced by Kellogg (1996). The primary suggestion of this writing model was that WM, which functioned as a short-term store and processor, tried to orchestrate many demanding activities such as planning, translating, typing, and monitoring ideas.

⁸ He noted that this respond can be shaped by the persona of the source text.

The main components of Kellogg's (1996) model are formulation, execution, and monitoring. The formulation includes two subcomponents. The first one is planning, which is about setting goals, thinking about ideas, and organizing them. The second one is translating, which is about converting (building a syntactic frame for some selected lexical units, etc.) and expressing ideas in a written form. The execution of these translated ideas requires programming and executing, which are related to the usage of the motor system depending on the output mode (handwriting, typing, etc.). Lastly, monitoring also includes two subcomponents. The first one is reading (comprehension, the establishment of coherence, etc.), which is considered a necessary but insufficient condition for writing well. And the second one is editing, which results from the comparison and mismatch between a writer's intention and the produced material. It can be global (paragraph organization, etc.) or local (spelling, etc.).

Kellogg (1996) also notes that the model does not imply straight progress from formulation to execution and monitoring. Rather, it implies a simultaneous (and nonlinear) activation of these processes depending on the capacity of the central executive. For example, either execution or renewed formulation may follow a monitoring/editing activity. Again, if execution is automatized enough, it can occur simultaneously as a writer formulates or monitors the written material.

These activities, particularly formulation and monitoring, are mentally effortful, and can be more demanding on WM resources: central executive, phonological loop, and visuospatial sketchpad (Table 1).

Table 1. The Use of Working Memory Resources by Six Basic Processes of Writing

Main Components	Basic Processes	Working Memory Resource		
		Visual Spatial Sketchpad	Central Executive	Phonological Loop
Formulation	Planning	X	X	
	Translating		X	X
Execution	Programming		X	
	Executing			
Monitoring	Reading		X	X
	Editing		X	

Note: Adapted from Kellogg (1996, p. 63).

For example, planning requires both the visuospatial sketchpad and the central executive. This is because it does not only include activities such as visualizing ideas or organizational schemes, etc., but also more demanding activities such as generating ideas or trying to establish an appropriate tone for a particular community (audience). On the other hand, translating requires a central executive together with a phonological loop because the phonological words selected in a syntactic frame are kept in this short-term store. Again, programming, an execution process, can demand central executive resource depending on previous writing practice (using motor output schemas, etc.), but executing do not⁹. Lastly, it is evident that reading to monitor the written passage requires the phonological loop and central executive. Still, the editing activity can demand and probably consume more central executive resources depending on the type of editing (detection of a spelling error vs. revision of passage organization, etc.)¹⁰.

⁹ This is one of the differences between Hayes (1996) and Kellogg (1996).

¹⁰ Referring to a study by Brown, McDonald, Brown, & Carr (1988), Kellogg (1996) reported that because of the limited resource of the central executive, the demands for an activity may be reduced or increased by the level of demands for another activity. For example, monitoring activities suffered and revealed more failures when the demands for the formulation increased by asking the writers to recall sentences (rather than reading and transcribing). Additionally, according to Kellogg (1996), the usage of central executive processes may depend on the automatization of semantic and phonological subprocesses.

Later, Kellogg, Whiteford, Turner, Cahill, and Mertens (2013) evaluated this model on the basis of experimental studies and suggested that there can be some adaptations, such as the possible inclusion of the phonological loop in editing (reviewing the text, etc.) and the involvement of the visual-spatial sketchpad with only concrete concepts. Beyond these, he recognized new WM models that do not separate storage components and define WM as attended representations, etc. (e.g., Cowan, 2011). However, he also noted that even if this is the case because both WM approaches embed a limited central executive into their models, they do not violate the basic assumption of Kellogg's (1996) model.

Hayes' (1996) model was also evaluated and revised many times (e.g., Chenoweth & Hayes, 2001; Hayes, 2012). Hayes and Berninger (2014) introduced one of the most recent revisions¹¹. Its framework consists of three interacting levels, resource, process, and control (Fig. 5). I will give more details about this model by focusing on the revisions in the following paragraphs.

¹¹ There were different versions of this model (e.g., Leijten, Van Waes, Schriver, & Hayes, 2014). Because this model was one of the most recent ones and was explanatory enough for this study's context, I preferred reporting it.

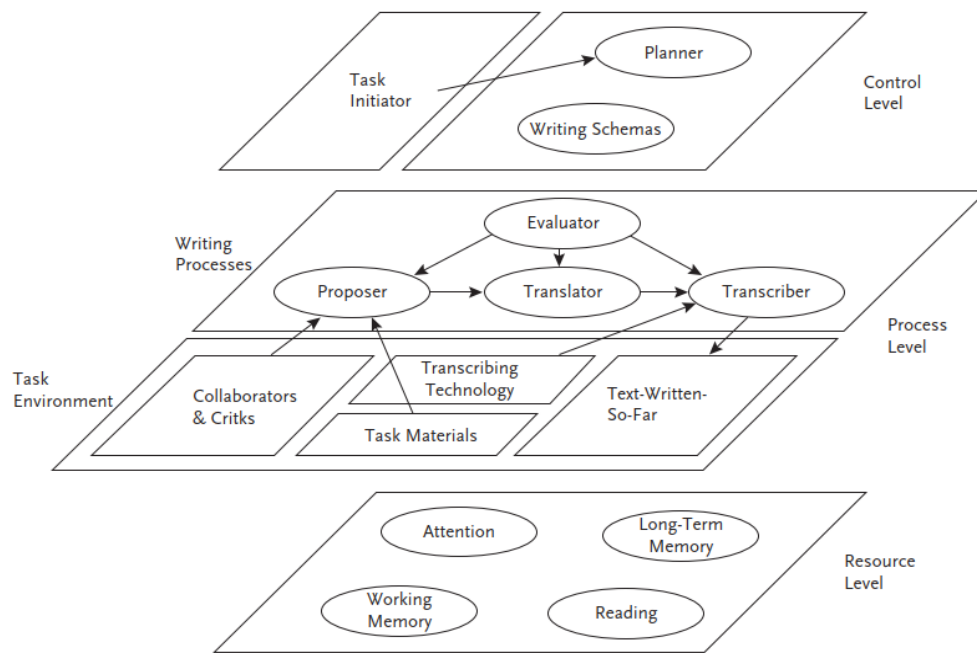


Fig. 5 The revised cognitive writing processes model

Taken from Hayes and Berninger (2014, p. 5).

An important addition to this revision is the attention component, which is also referred to as executive control/function and associated with WM and which is defined as the ability to focus on a task in the presence of distractors. Thus, the resource level includes attention, WM¹², LTM (audience knowledge, etc.), and reading. According to the model, these components can interact with each other, as well as with processes from other levels. For example, attention resources can be used by top-level control processes such as the task initiator.

The process level includes task environment and writing processes planes. There is no essential change with respect to the task environment, consisting of collaborators & critics (immediately present ones such as teachers), transcribing technology, the text

¹² Unlike his 1996 model, he does not refer to Baddeley's (1986) model in this revision.

written so far, and task materials. When it comes to writing processes, it includes a proposer, translator, evaluator, and transcriber¹³. The proposer's function is to get input from the planner, the text written so far, LTM, or the task environment, and to suggest nonverbal ideas for the text. The translator gets these ideas and converts them into verbal representations. It can also get verbal ideas from verbal LTM and convert them into new language strings (Hayes & Chenoweth, 2007), and its fluency is said to be dependent on available WM resources, as well as the linguistic experience of a writer. The third component of the writing process is the transcriber, which turns the grammatical strings taken from the translator into text. The last writing process, evaluator, can judge and ask for the revision of the outputs passed by any other processes. For example, it can reject a proposed idea before its translation, etc.

Lastly, the control level consisted of task initiator, planner, and writing schemas. The task initiator is simply the person who starts the writing task (e.g., a teacher, the writer herself, etc.). It will probably specify the topic, the audience, etc., and influence the planner which sets writing goals. The range of these goals may be dependent on the level of experience of a writer. For example, novice writers can focus on a single topic, but experienced writers can order some topics and subtopics, and they can consider the tone and its impact on the audience. Finally, writing schemas include writers' beliefs about the target properties of a text (genre knowledge), as well as how to produce it (strategic knowledge). Again, just as the function of a writing planner, writing schemas can be dependent on the writer's experience or preferences. For example, a writer may adopt a kind of stream-of-consciousness writing strategy, which may change the

¹³ These were adapted from an earlier model of written production (Chenoweth & Hayes, 2001)

selection of writing processes and the interaction pattern between them (e.g., the proposer does not need to keep a general topic in focus, etc.).

Another relatively recent model which quite explicitly referred to the role of EF and audience was introduced by Kellogg (2008). Before giving details about this model, I need to consider a relevant developmental writing model which introduces some key terms such as knowledge telling and knowledge transforming. This model was introduced by Bereiter and Scardamalia (1987). The suggestion was that writing development could be relatively independent of other language skills, such as speaking during which an audience is available and can follow a relatively different developmental path.

According to this model, novice writers¹⁴ tend to use a knowledge-telling strategy to directly use readily stored/organized information to reduce the demands on limited cognitive resources. Children reach proficiency in speaking at an early age, but catching the same proficiency in writing takes time. One of the most apparent reasons for this asymmetry is the difficulty in automatizing writing code, but there is much more than that. A less obvious but essential reason/obstacle is the absence of an audience in writing. Content generation is not a big deal in a conversation mainly because conversational partners support the speaker in many ways. However, without this support, children can have difficulties keeping track of a topic and making appropriate choices to organize the passage. Thus, it can be difficult for them to make the passage more understandable for the absent audience. To deal with these kinds of challenges, novice writers discover alternative cues to retrieve information from memory (Bereiter

¹⁴ They prefer using the words mature and immature, but I prefer using the words novice and expert to be consistent with the literature.

& Scardamalia, 1987). One of these cues is the already-produced text. Although it is not enough to organize the text coherently, this text can create a kind of stream of consciousness. Related to this, the other cue can be the topic assigned by a teacher. It can help the novice writer keep track of the content. The last cue is the previously stored genre-specific elements and their organization (narrative, argumentation, etc.). Based on this argument, Bereiter and Scardamalia (1987) suggest that novice writers use a knowledge-telling model of discourse production. They use the abovementioned cues and retrieve readily available knowledge to deal with problems such as the absence of external support. In other words, the written output produced by novice writers is dependent not only on language abilities but also, for example, on their topic knowledge (Bereiter & Scardamalia, 1987).

On the other hand, more experienced writers adopt a knowledge-transforming model (Bereiter & Scardamalia, 1987; Scardamalia & Bereiter, 1987). This complex problem-solving model embeds knowledge-telling as a subprocess and considers knowledge and operations (e.g., using knowledge to infer another knowledge, etc.) in content and rhetorical spaces. In the content space, the knowledge states include beliefs, and operations include transforming these beliefs (inferring, hypothesizing, etc.). However, in the rhetorical space, the knowledge states include rhetorical context representations such as the written text and goals, and operations include the changes in text, goals, or their relations. What differentiates this model from the knowledge-telling model is that expert writers do not only use previously stored knowledge but also continuously detect rhetorical and content-related problems and create solutions to them.

When a decision is made to deal with one of these problems, it will have implications for the others.

Later, Kellogg (2008) introduced a model which considered the knowledge-telling and knowledge transformation stages but also focused on a third, knowledge crafting stage. Notably, he considered the importance of the audience and associated the successful consideration of the audience with expertise. He also emphasized that EF capacity is the primary constraint on writing development across these stages, and reducing the demands on this system (e.g., by automatizing language production) can help a writer to achieve the necessary control. He suggests that text composition is a type of what-to-say (content) and how-to-say (rhetoric) problem, which consumes the limited WM and attention resources. And just like in other kinds of learning, such as playing chess, it takes time to mature and gain executive control of cognitive processes to respond to the requirements of a task. Actually, according to him, it takes decades to gain expertise in writing: at least two decades of maturation and practice to move from simply telling what a writer knows (knowledge telling) to transforming what is known for the benefit of an author (knowledge transformation), and finally, to the crafting of what is known for the benefit of the reader.

According to him, at the most advanced knowledge crafting stage, the writer can simultaneously maintain three components: the writer's planned content, the text itself, and the possible interpretation of imagined readers. Thus, for an expert writer, executive attention is not allocated only to basic writing processes such as language production or planning but also to these three alternative content representations that can contradict each other. He emphasized that, for example, reader representation can also be available

at an early stage, but its maintenance in WM may be unstable: the audience can guide the word choices but not the review of a text, etc. Therefore, being aware of a reader's perspective can be a critical/necessary condition, but it is not sufficient to read and interpret the text from the perspective of a fictional reader; enough WM system resources must be available to coordinate the interaction among the author, the text, and the reader in a relatively easy way.

More recently, Kim and her colleagues also contributed to the literature by introducing, testing, reevaluating, and revising a comprehensive hierarchical model, the DIEW, which quite explicitly and systematically considers EF resources and audience processing (or reader PT).

Initially, the model was introduced and tested on native English speakers (Kim & Schatschneider, 2017). Later, it was extended to Korean speakers by Kim and Park in 2019. And more recently, Kim and Graham (2021) worked on it and created a more comprehensive version that specifies the directions between the processes more explicitly.

DIEW is developed over an earlier model which differentiated lower-level language (grammatical knowledge, etc.) and cognitive skills (WM, etc.) and higher-level cognitive skills (inference making, etc.) (Kim, 2016). The essential suggestion of the model is that the writing activity needs to generate ideas, translate them into oral language and then transcribe them into written text (Kim & Park, 2019). And as this happens, many component skills which have hierarchical structural relations are involved (Kim & Park, 2019) (Fig. 6).

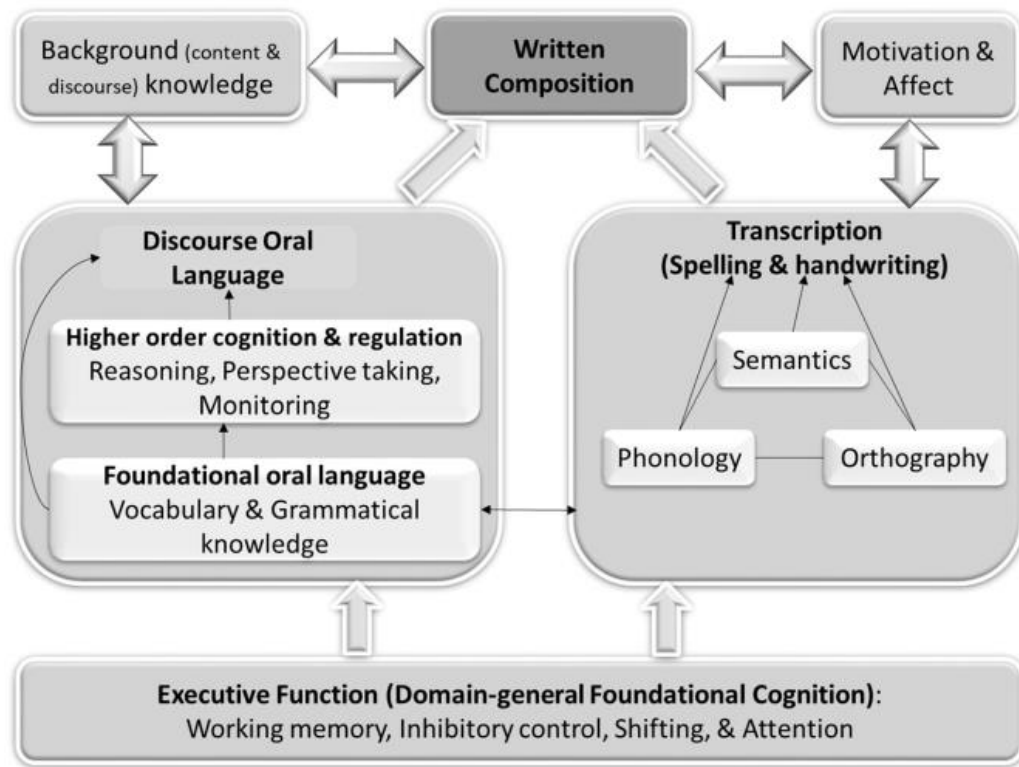


Fig. 6 Direct and indirect effects model of writing

Taken from Kim and Park (2019, p. 1320).

At the lowest level, EF (e.g., WM and AC) establishes a necessary domain-general foundation for basic language skills, such as vocabulary and grammar knowledge, which are essential requirements for a writing activity.

The effect of EF on writing can be mediated by discourse oral language, and transcription skills. The discourse oral language, which is defined as ‘the ability to produce oral texts’ (p. 1321), is suggested to contribute to text production by generating, organizing, and translating ideas. To be able to do this, it makes use of other skills from different levels. First of all, it requires EF resources. Additionally, it draws on low-level oral language skills such as vocabulary and grammatical knowledge and higher-level cognitive skills such as inference, reasoning, PT, and monitoring. Thus, for example, PT

can require EF, and writing can require PT, and this can result in partial mediation of the EF's effect on writing. As reported above, transcription skills are the other possible mediator between EF and writing. Particularly, spelling and handwriting, which require phonologic, orthographic, and semantics knowledge and awareness, can mediate the effect of EF. It is reported that the extent of these transcription skills' automatization can determine the use of mental resources (WM and AC). If they are not developed enough, the employment of other component skills during writing activity can be constrained.

Apart from these, it is stated that writing would be influenced by background knowledge, including discourse and content knowledge. The first, discourse knowledge, contains genre, and a writer needs to know both the characteristics of a genre (text structure, etc.) and the way to present it (procedures and strategies). The second type of background knowledge, content, can be essential particularly for novice writers who tend to directly report their knowledge. The development of this knowledge is expected to correlate with writing (and oral) language development. Lastly, the model recognizes the influence of motivation and affect (e.g., anxiety) on writing (please see Kim & Park, 2019 for detailed literature support for all these components¹⁵).

One contribution of this model is emphasizing the importance of high-level skills such as PT, defined as “one's knowledge of and inferences about others' mental and emotional states” (Kim & Park, 2019, p. 1323). It is suggested that audience awareness, which is a kind of PT¹⁶, is essential (e.g., Carvalho, 2002; Wollman-Bonilla, 2001)

¹⁵ The hypothesis of the model was tested both for English (Kim & Schatschneider, 2017) and Korean-speaking children (Kim & Park, 2019), and both studies revealed that DIEW fits the data and higher-level processes mediate the contribution of lower-level processes.

¹⁶ They used PT as a broader concept, suggesting it goes beyond audience awareness.

because writers have to consider readers' needs and adapt the text accordingly (Kim & Schatschneider, 2017). Another thing DIEW emphasizes is that the involvement of these higher-order processes increases with the automatization of lower-level processes, such as transcription. This is because these low-level processes may dominate and consume particularly a novice writers' limited EF resources, which creates the basis for all these processes.

Later, Kim and Graham (2021) evaluated Kim and Park's (2019) DIEW model, talked about its advantages, introduced a further expanded DIEW, and presented another demonstration of it. In this introduction, they suggested that DIEW articulates three related hypotheses, hierarchical, interactive, and dynamic, about the structural relations of component skills. First, the hierarchical hypothesis specifies some pathways that relate component skills and lower-level cognitive skills to written composition. According to the model, these relations are not always direct. On the basis of previous findings which indicate that WM is crucial for writing performance (e.g., Hayes & Chenoweth, 2007), for component skills such as grammar and vocabulary (e.g., Kim, 2017), and for higher-level cognition such as discourse oral language (Kim, 2016) which is associated with written composition (Kim & Schatschneider, 2017), the model suggests that the contribution of WM on writing can be, for example, via grammatical knowledge and discourse oral language. Secondly, the interactive hypothesis suggests that, as they develop, component skills interact with each other and with writing overall (Kim & Park, 2019). Lastly, the dynamic relations hypothesis suggests that the weight of associations between component skills and written composition changes with development. For example, as Kim and Park (2019) indicated, in the beginning,

transcription processes/skills can determine writing performance. But, with the development, discourse oral language can gain more importance (Kim & Graham, 2021).

After emphasizing these hypotheses, Kim and Graham (2021) extended DIEW by particularly including the reading skill and emphasizing the importance of considering written product dimensions (Fig. 7).

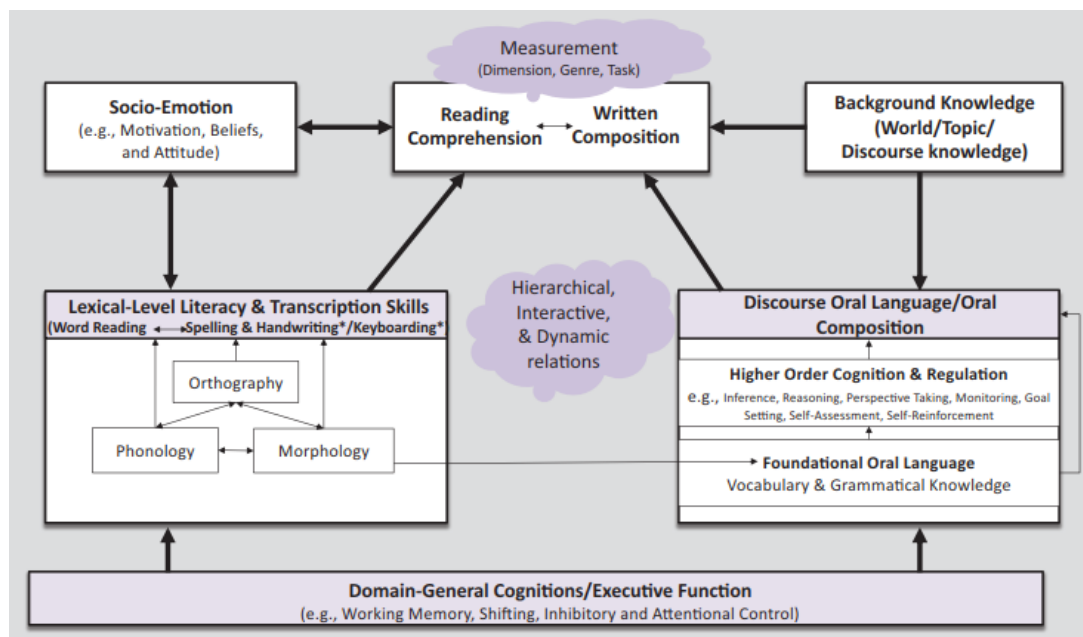


Fig. 7 Expanded direct and indirect effects model of writing

Taken from Kim and Graham (2021, p. 2).

The basic rationale behind the inclusion of reading skills and their interaction with writing is that a writer needs to read and interpret her text as she revises it. Furthermore, a writer frequently needs to read and analyze source texts to create some knowledge basis for her writings. To be able to do this, at least word reading is necessary. Otherwise, a writer cannot decode the words that she wrote, and she cannot

create an accurate mental representation to compare with her writing goals, which may result in low-quality products. They also suggest that the pattern of the contribution of reading can be dependent on developmental stages: for example, at initial stages, lexical-level literacy skills can partially mediate the influence of component skills on discourse oral language, but with reading development, the influence of these basic skills can be less influential and reading comprehension can start completely mediating the influence of oral language on writing quality.

The second important reason for modifying DIEW was the dynamic/differential relations hypothesis (Kim & Graham, 2021). The basic idea behind this hypothesis was that a written output could be approached from different dimensions (writing quality¹⁷, writing productivity, accuracy, syntactic complexity, etc.), and the effect of high-level processes (inference, monitoring, PT, etc.) on these dimensions are not probably fixed. In other words, DIEW suggests that if there are various dimensions of a written product, then the relations of component skills to these dimensions can differ, and for this reason, a theoretical model should consider both outcome (dimensions) and predictors (component skills). For example, suppose the focused dimension is the quality of the ideas. In that case, the ability to use vocabulary and grammar knowledge, the ability to consider audience needs, and the ability to arrange the passage coherently, as well as the ability to transcribe well, must be important (Kim & Graham, 2021). However, if the focus is on writing productivity (the length of the passage), transcription skills rather than higher-level skills will probably be important. Again, if the focus is on accuracy, spelling ability and grammatical knowledge can be more critical. Previous empirical

¹⁷ The quality and coherence of the content, as well as language usage.

studies revealed data that support this approach. Higher-order cognition, such as PT, is associated with the quality of a written passage (Kim, 2020; Kim & Graham, 2021) but not with writing productivity and accuracy (Kim & Graham, 2021). Reporting these, Kim and Graham (2021) say that previous models were silent about the importance of and the implications of dimensionality. According to them, this hypothesis must be part of a writing model because the involvement of component skills will probably vary depending on the focused dimension(s).

Having reported these models, it may be important to note that their primary focus was on native speakers. However, studies indicated that the component skills, writing processes, or cognitive processes introduced in these models are also influential in L2 context (e.g., Graham & Eslami, 2020).

For instance, Silva (1993) reviewed 72 studies and concluded that there are some quantitative differences with respect to the involvement of writing processes and components, but qualitatively similar processes are included in both first language (L1) and L2 writing (Silva, 1993). As an illustration of the quantitative difference reported by Silva (1993), writing in L2 was relatively more difficult and less effective (lower holistic scores) than writing in L1. Regarding language usage, L2 writers' sentence structures were simpler, T-units were shorter, clauses were longer but fewer, coordination was more, but subordination was less. Overall, the texts were less accurate and less fluent. Regarding the discourse, the responses to argumentation tasks, etc., were less effective (less defining, exemplifying, etc.), and the reader orientation was less appropriate (e.g., underestimation of reader knowledge). There were different patterns in the usage of cohesive devices; for example, L2 writers used more conjunctive but less

lexical cohesive devices. And their overall lexical control, variety, etc., were less effective. Regarding writing processes, L2 writers planned less and had difficulty in goal setting, generating, organizing, and transcribing materials (less fluent and productive). Although they also had difficulty in revising, they revised more than L1 writers; however, they reviewed, reread, and reflected on their texts less.

Another important support for the existence of overall qualitative similarity between L1 and L2 processes was provided by Kim, Tian, and Crossley (2021). They tested Hayes and Berninger's (2014) writing model in L2 undergraduate students and revealed that similar processes are involved in L2 writing: There was an association among cognitive and linguistic resources, writing processes, and the quality/length of written products. Hence, although there were some quantitative differences, the overall writing processes of L1 and L2 writers were qualitatively similar to each other, and the suggestions of the models (e.g., DIEW) can be extended to L2.

CHAPTER 3

DIRECTING ATTENTION TO AUDIENCE

In this chapter, I will narrow my focus to three components that can influence the employment of EF resources and affect the performance of an EFL writer. The first one is the language proficiency of an EFL writer. Although I will not directly study the influence of L2 language proficiency as a variable in this research, it may be essential to create a background for why I choose a relatively lower proficiency group (upper-intermediate level learners), and how it can interact with EF. Secondly, I will focus on the audience and try conceptualizing highly associated audience awareness and PT concepts. I will also consider the recruitment of the limited cognitive resources (EF) by this high-level PT and low level non-automatized language processes. Lastly, I will concentrate on the possible role of an audience-oriented task that can direct the writer's attention to the audience, help novice EFL writers manage their EF resources, and increase writing performance.

3.1 Language processing

In this study, I will focus on relatively lower language proficiency (~upper-intermediate) participants, thinking that the efficiency of my manipulation can be shown by this group of EFL learners better.

As would be expected, although it may not be sufficient, high L2 proficiency is necessary for composing well-written L2 passages (Cumming, 1989; Crossley & McNamara, 2012; Ito, 2004). This is because, for example, L2 writers need to transform

nonverbal messages, such as propositional content/ideas or audience awareness, into appropriate verbal-linguistic forms, and if L2 writers' language proficiency is not high enough, they can have difficulties, for example, in connecting parts of a text appropriately, which can make L2 writing more challenging (Schoonen et al., 2011). Moreover, less proficient writers tend to create L2 sentences, for example, by generating ideas in their L1 and then translating them to L2 (Wang & Wen, 2002). This dependency can also increase writing difficulty and may be the other factor behind the low text quality of the passages written by EFL writers (Weijen, Van den Bergh, Rijlaarsdam, & Sanders, 2009).

The lack of enough language background may be one of the most important reasons for being unable to write efficient passages. However, as the writing models suggested (e.g., DIEW), it is not the only variable that influences, and its presence does not guarantee the success of writing. Essential for this study's concern, it may also be essential to consider, for example, the involvement of one of the critical cognitive components, EF.

An important consequence of having low language proficiency may be more recruitment of the limited EF resources. As reported above, it is widely accepted that EF resources, such as WM, create a foundation for writing (e.g., planning, transcribing, and reviewing) (e.g., Kellogg, 2001; Kim & Park, 2019), which includes basic language processes (e.g., Weigle, 2005). This may be because writing processes, such as turning the retrieved and processed ideas into written forms (Kim, Tian, & Crossley, 2021), recruit language knowledge (e.g., handwriting, vocabulary, and grammar), and this

recruitment process may require EF resources particularly if the language proficiency level of an EFL learner is not so high (Serafini & Sanz, 2016).

An important concept that can help us understand the reason behind this association and which is related to EF and L2 proficiency may be automatization. Some researchers operationalize it as a kind of fast, ballistic (unstoppable), effortless (without cognitive load), and unconscious information processing, while some others, at least implicitly, assume the existence of these characteristics but focus on the underlying processes of it (e.g., what/how something is automatized, the role of automatization in skill acquisition, etc.) (Segalowitz, 2003). DeKeyser (2001) is one of the researchers who adopted the latter approach and, based on previous findings, suggested that with automatization, there can be less memory load (capacity-free processing), and this relative independence from the memory system can be associated with faster and more accurate processing. Importantly, this capacity-free suggestion is in accordance with the definition, which suggests that WM is involved more, particularly if the target processes are not fully automated. In other words, if a process is not automatized enough, it will require conscious processing, hence WM (EF) recourses (Baddeley, 1986).

Turning back to the association between language proficiency and the usage of EF resources, the more the writers practice their basic language processes (e.g., via sentence combining), the more automatically they can process (Graham & Perin, 2007) and retrieve these kinds of basic information from LTM and automatically check, for instance, the subject-verb agreement, etc. (Kellogg, 2008). On the other hand, particularly if the writer's language proficiency, such as in encoding and retrieving syntactic structures (Vasylets & Marin, 2021) and associating orthography with the

lexical entry (McCutchen, Covill, Hoyne, & Mildes, 1994), is not high (automatized) enough (Linck, Osthus, Koeth, & Bunting, 2014), writing activity can demand more resources (e.g., for attending to L2 language and suppressing L1 cues) (Kormos, 2012; Sunderman & Kroll, 2009).

3.2 Audience processing

Because their audience is generally distant readers who cannot respond to their passages and are not personally known by them, “writers must abstract their potential readers, at least to a certain degree” (Magnifico, 2010, p.170). Importantly, this can be the case even when a passage is written to an active/participatory audience whose probable reactions are known. There can be two sources of this controlled abstraction/imagination (Gibson, 1950; Magnifico, 2010). Firstly, it can be based on the generalizations from the writer’s own preferences, characteristics, etc. Secondly, it can be based on readers’ (external) features (e.g., the general content of the magazine they read). The latter is associated with good writers who take the reader’s perspective and consider the audience’s needs to adapt their passages.

One important concept to consider in this context is PT. It is defined as “the active cognitive process of imagining the world from another’s vantage point or imagining oneself in another’s shoes to understand their visual viewpoint, thoughts, motivations, intentions, and/or emotions” (Ku, Wang, & Galinsky, 2015, p. 94-95). Importantly, a successful PT, which can engage the cognitive processes and make others’ perspectives more cognitively accessible (Ku et al., 2015), can increase the overlap between the mental representations of the self and the other and, in this way,

help to bond with others more effectively (Galinsky, Ku, & Wang, 2005). Some researchers differentiate and operationalize PT as a trait versus state PT. If it is a trait of personality, it is about the nature of individuals, and it is either given by nature or developed as a result of developmental experiences (Duan & Hill, 1996). On the other hand, state PT is situation-specific (Duan & Hill, 1996). Its appearance may depend on situational factors¹⁸ relatively regardless of the personality characteristics. Still, the availability of high trait PT tendency may make state PT easier. This is because, for example, those who have higher PT tendency can judge other people's unique level of traits (Colman et al., 2017) and take their perspectives in different conditions better (Crocetti et al., 2016).

According to Cho et al. (2021), the operationalization of PT is also associated with theory of mind, epistemological understanding, and audience awareness. The theory of mind, which develops in early childhood, is defined as the ability to comprehend the mental states of others (what they know, want, feel, believe) and to predict their behaviors (Baron-Cohen, Leslie, & Frith, 1985). When it comes to epistemological understanding, it is about the development of PT with respect to someone's approach to knowledge (Cho et al., 2021). This development continues during adulthood and consists of four stages (Kuhn, Cheney, & Weinstock, 2000). At the initial (realist) stage, a person takes a single perspective and accepts the knowledge as an objective truth without critically evaluating it. However, at the final (evaluativist) stage, the knowledge is considered a constructed phenomenon and is critically questioned/evaluated from different perspectives. Importantly, Kuhn et al. (2000)

¹⁸ For example, an audience-oriented task which encourages a writer to take reader perspective would probably require state PT.

compared different age groups and revealed that almost all undergraduate students could take multiple perspectives while evaluating knowledge. The last concept associated with PT is audience awareness¹⁹ which is defined as considering the audience's perspective (their knowledge, needs, etc.) and trying to adapt a passage accordingly (Cho et al., 2021).

On the basis of this short review, since this study will focus on adult preparatory school university students, I assume that their theory of mind is developed, and although their ability can range on a continuum, they have the ability to take the other people's perspectives²⁰.

It may be important to report once again that the models I summarized above recognize and emphasize the importance of considering the audience in writing. And particularly a more recent one, DIEW, directly includes PT as an important writing component. According to this model (Kim & Park, 2019), PT, which is a broader concept than audience awareness (e.g., considering the reader needs during writing), contributes meaning making in many different ways. As Cho et al. (2021) reported, it may add the depth and coherence of a passage, for example, by helping the writer to understand the aim and expectation of a writing task and to communicate more effectively. Again, it may help using information resources (to be used in the written passages) more critically, for example, by considering the source text's author's motivation for writing and by evaluating his/her credibility. Furthermore, and most

¹⁹ Also known as sense of audience.

²⁰ I will also consider the literature that associates PT with the theory of mind or being able to take multiple perspectives, but instead of using these terms, I will use the term PT throughout this chapter to be consistent.

importantly, it may support effective writing by helping to consider the audience (e.g., putting the writers into the shoes of their audience) and to adapt the passages accordingly. Hence, better PT may be associated with better passages and, as Cho et al. (2021) and many researchers in the field suggested, this may be mainly because the writers must be aware of their audience (Bereiter & Scardamalia, 1987; Carvalho, 2002; Hayes, 2012) and try to consider their needs, for example, their background knowledge and attitudes towards the topic (Weigle, 2005) as well as their linguistic and cognitive competence (Roxßnagel, 2000) in order to adapt and increase the efficiency of the content.

However, taking other's perspective in a specific condition, state PT, can be a cognitively demanding activity that requires the usage of EF resources, such as WM (e.g., Kim, 2020; Kim & Graham, 2021; Ryskin et al., 2015; Schneider, Lam, Bayliss, & Dux, 2012; Wardlow, 2013) and AC (e.g., inhibition of a distractor)²¹ (Brown-Schmidt, 2009). Hence, an incident which leads to cognitive load²² may result in impairment in PT (Roxßnagel, 2000), let alone adapting a passage accordingly. This can explain the findings, which show that the reader support in a passage can be determined by WM capacity (Alamargot, Caporossi, Chesnet, & Ros, 2011). Thus, the higher the EF capacity, the more resources can be allocated to writing processes, including taking the reader's perspective.

The associations between limited cognitive resources, linguistic processes, and PT can be generalized to L2 writing. L2 writings by undergraduate students revealed

²¹ For example, a person who needs to take another person's perspective may need to inhibit her own (egocentric interpretation) while maintaining the other's perspective.

²² It is defined as the need to allocate limited cognitive resources for simultaneous processes.

that cognitive (e.g., WM and AC) resources predicted the efficiency of the linguistic processes (e.g., vocabulary knowledge), which in turn predicted the quality of the written products²³ (Kim, Tian, & Crossley, 2021). As reported above, another process predicted by WM is PT (Roxßnagel, 2000), and importantly, this higher-order cognitive skill can also predict the writing quality both in L1 and in L2 (Kim, Wolters, Mercado, & Quinn, 2021). In other words, linguistic and PT performance can predict high-quality L2 passages (e.g., having a well-structured organization with a logical flow of detailed ideas). However, because both processes require EF resources, this dependency on the same limited resource can create competition, which may impair PT (Roxßnagel, 2000), and reduce audience processing (Carvalho, 2002) because of the priority given to the fundamental language processes (Kim, 2020; Kim & Graham, 2021).

3.3 Designing a task directing the EFL writer's attention to the audience

An important problem in EFL writing may be the allocation of less EF resources for high-level audience processing, which may result from several reasons, such as the dominance of not sufficiently automatized language processes (Kim, 2020; Kim & Graham, 2021), the lack of enough EF resources (McCutchen, 1996), or simply (as a novice writer) not being aware of the importance of audience who is not immediately available (Traxler & Gernsbacher, 1993). The question is whether designing a task that

²³ There can surely be some cases that indicate well-automatized language processes but not a good writing pattern. I do not suggest that automatized language processes directly lead to well-written passages, what I suggest is that language processes are essential in writing, and if not developed enough, they can interrupt higher-level writing processes such as taking the reader's perspective and adapting the passage accordingly. In this study, I will consider relatively low (upper-intermediate) level L2 learners to be sure of the existence of competition for using the limited EF resources.

directs those (e.g., upper-intermediate level) EFL writers' attention to the audience may help them keep the audience in their minds and manage the allocation of EF resources more efficiently, for example, by adapting their passages while considering the audience/reader perspective, which may be reflected in their increased writing performance.

As stated above, PT ability can be essential to consider audience and create high-quality passages. However, although an average adult can take another person's perspective, this does not mean that PT (or mind-reading inference) is automatic under every circumstance (Apperly, 2018). Empirical studies confirmed this suggestion. For example, Apperly et al. (2006) indicated that the (other's perspective) false belief²⁴ of a character was processed only when it was relevant to the task. In this study, the participants were told that they were going to watch a video and were assigned one of the three instruction conditions. In one of these conditions, the participants were not explicitly instructed. In the second condition, they were asked to attend to an object's (final) real-location after the change. And in the last condition, they were asked to attend to both the real location and the woman's belief about the location. After the instructions, they watched the video in which there were a woman and two boxes. The woman in the video gave a clue about the location of an object which was put into a box, and she left. Then, another man arrived and changed the location of the boxes. Thus, the participants who watched this video need to update the location information in their minds and must be aware that the woman has a false belief. After the video was

²⁴ For example, suppose that you were together with a person, but then something changed in her absence. Now, you are aware of the change, but she still believes that the environment is the same. This is her false belief.

stopped, to understand whether the participants automatically took the woman's perspective, the researchers asked the participants to answer one of two probes. The first one was a reality location probe, “It's true that it's in the box on the right [left]”, and the second one was a false belief probe, “She thinks that it's in the box on the left [right]” (p. 842). The results revealed that participants were slower to make judgments about the woman's belief when the task did not give instructions and directed their attention to taking her perspective. Thus, PT, which is a kind of mind-reading inference, may not be automatic, and explicit instructions can make a difference in adults (Apperly, 2018). In other words, even if a person has PT ability²⁵, the content of a task may be needed to direct attention of him/her to use this ability efficiently.

Whether the direction of attention resources through task instructions improves writing performance is also considered in the writing literature. Limited Attentional Capacity model/Trade-off Hypothesis (Skehan, 2009) suggested that WM capacity is limited, and because of this, the direction of the attention through a task can result in a trade-off between different dimensions; that is, there can be an improvement on only one dimension. For example, if a task directs the writer's attention by manipulating task complexity, the task-based improvement on the final product can be either in its complexity or accuracy. However, Cognition Hypothesis (Robinson, 2007) suggests that the direction of attention resources through the task instructions²⁶ can increase the mental effort (the level of notice) of the writer, can help her to make cognitive

²⁵ Apperly, Simpson, Riggs and Chiavarino (2006) did not assess the trait PT in the study. However, thinking of the random assignment of participants into three groups, I assume that the trait PT ability was similar across the groups.

²⁶ For example, by making multiple PT a part of the task and increasing its complexity in this way.

distinctions in her language processes, and can improve her writing without a trade-off, for example, between accuracy and complexity (Robinson, 2007; Robinson & Gilabert, 2007). Many studies have been conducted, and both models are supported to some extent. This study aims not to compare these models, but they were reported to show the potential of a task to direct the attention resources and to improve writing performance.

This can be the case because the direction of attention through instructions can help participants take other people's perspectives (Apperly et al., 2006), and this can encourage writers to try to take a reader's perspective and increase the writers' awareness of the audience's needs (Hollaway, 2004; Hollaway & McCutchen, 2004), which in turn can increase their writing performance. In accordance with this suggestion, some studies indicated that increasing audience awareness through a writing task can improve the writers' performance. For example, Traxler and Gernsbacher (1993) conducted a pre-post-designed experiment to investigate the effect of a PT intervention. In the study, L1 English undergraduate students wrote a passage describing geometric figures. They had been instructed that another person would read their texts and, based on their descriptions, try to select the appropriate figure among the similar-looking distractors. Then, these students were divided into two groups. The PT group was shown new figures and instructed to use the written descriptions (by other participants) in order to select the appropriate figure. In contrast, a control group simply ranked the figures on the basis of some instructions. After this intervention stage, all participants revised their original passages in another session. The evaluation of the revised passages revealed that the PT group improved their texts more than the control group. The suggestion was that the experimental group had an opportunity to represent

the readers more accurately after using the texts written by others to choose some figures. In other words, they improved their awareness of how the audience could perceive their passages and revised their texts accordingly.

Another study was conducted by Cohen and Riel (1989) to compare the essays written for midterm evaluation (teacher) with the essays written for overseas communication (peers in other countries). The analyses revealed that writing for the latter (distant audience) resulted in higher performance: better language use, organization, content (e.g., explicitly expressed ideas), etc. Again, Block and Strachan (2019) conducted a study and indicated that writing for an external audience (e.g., a best friend or a librarian who is outside the classroom), compared to writing for an internal audience (e.g., classmates and teacher), produced higher quality texts. Lastly, Cho and Choi (2018) compared two writing prompts, one of which specified an audience who did not have knowledge about the topic, while the other one asked to write for evaluation purposes²⁷. The results indicated that when the prompt made writers aware of the audience more, they tried more to adapt their texts to respond to the readers' needs (e.g., introducing a context).

Although their instructions did not explicitly state, in one way or another, all of the above-reported tasks directed the writer's attention to the target audience's perspective. For example, when I think about the study by Cho and Choi (2018), the participants needed to take the perspective of a reader who did not have background knowledge about the topic. Otherwise, it would not be possible to organize their

²⁷ To note, one suggestion of these kinds of studies is that if no audience was mentioned in a writing prompt, there would be an implicit assumption that the aim of writing is not communication but evaluation (Cho & Choi, 2018).

passages accordingly. Hence, these studies indicate that even a single writing prompt can increase the writers' audience awareness and orient the writer to the audience.

Based on the literature reviewed so far, I suggest that even relatively low (e.g., upper-intermediate level) language proficiency writers can benefit from an audience-oriented task with instructions that explicitly increase audience awareness, for example, by encouraging them to take the audience's perspective. This can be the case because language processes, of particularly low-proficiency writers, may dominate and overload cognitive resources (Kim, 2020; Kim & Graham, 2021) which would be needed by audience processing (Ryskin et al., 2015; Wardlow, 2013). However, if the task processing instructions direct their attention to the audience, they can balance and manage their resources better. They can resist the pressure of competitive low level language processes, allocate some resources for keeping the audience in their minds while writing and adapting the text, and this can increase their overall writing performance.

Moreover, this kind of task may also have the potential to increase the level of persuasiveness of a passage. EFL university students are reported to have some difficulties in developing persuasive texts by providing the reader with reasons (Varghese & Abraham, 1998) and considering the alternative/opposing views (Qin & Karabacak, 2010), which may be associated with PT (Reznitskaya et al., 2001), and this is the case particularly if their proficiency level is relatively low (Cheng & Chen, 2009). This may be because, as I reported above, they may be prioritizing and allocating more of their EF resources for basic language processes. If this is the case, explicitly directing their attention to consider the audience may help those novice EFL writers. Doing this

may increase their awareness, for example, about the audience's needs and expectations, and they may support their opinions with more convincing reasons while also addressing potential counterarguments. This would be in accordance with the research, which indicated that thinking about the audience may potentially encourage writers to consider the other party's view and increase the quality of their persuasive passages (Nussbaum & Kardash, 2005; Wang, 2016). Thus, a task that directs the direction of the writer to the audience may increase not only the overall writing quality but also the level of persuasiveness.

CHAPTER 4

PRESENT STUDY

This study will consider the EF capability and trait PT tendency of upper-intermediate level novice Turkish EFL writers, direct their attention to the audience via an audience-oriented writing task and investigate its influence on the overall writing quality and the level of persuasiveness. The aim is to see whether the experimental writing task improves the writing performance by interacting with EF capability (low versus high) trait PT tendency (low versus high) of the participants.

4.1 The theoretical and empirical background

As mentioned in Chapter 2, many cognitive writing models emphasized the importance of EF as a core ingredient that orchestrates and monitors many demanding activities (e.g., Kellogg, 1996). They also recognized the importance of considering the audience for producing high-quality passages (e.g., Hayes & Berninger, 2014; Kellogg, 2008; Kim & Graham, 2021) and noticed that the efficient consideration of the audience might be associated with the availability of the limited EF resources (e.g., Kim & Park, 2019) which determines the writing performance (McCutchen, 1996).

Importantly, the availability of these resources may also be influenced by the automatization of fundamental linguistic processes. This is because conscious processing is dependent on EF (Baddeley, 1986), and the possibility of the involvement of higher-order processes, such as considering the audience in a writing task, may increase with the automatization of basic language processes (Kim & Graham, 2021). If

not automatized, these processes, such as connecting parts of a text appropriately (Schoonen et al., 2011), may overload or dominate the limited EF resources that a writer needs for taking audience perspective and creating efficient reader-oriented passages.

High EF capability may have the potential to reduce the effect of this limitation because, in that case, the possible load created by relatively less automatized language processes (Serafini & Sanz, 2016) and state PT [maintaining audience while writing] (Roxßnagel, 2000) can be managed relatively more easily. In other words, even if low-level language processes may dominate and consume an EFL learner's EF resources, an individual with relatively higher EF capability may spare some cognitive resources for considering the audience sufficiently (Alamargot et al., 2011). Still, because of the dominance of fundamental language processes (Kim, 2020; Kim & Graham, 2021; Schmidt, 1992) and difficulty of successfully integrating the audience into writing processes (Kellogg, 2008) or because of not being aware of the (importance of) audience who is not immediately available (Becker, 2006; Cho & Choi, 2018; Traxler & Gernsbacher, 1993), novice writers may ignore them (Hayes, 1980b).

Modifying a task with a part that directs the relatively low proficiency (upper-intermediate) novice EFL writers' attention to the audience may be one possible way to increase their awareness and help them manage their limited EF resources more efficiently. As Hayes and Berninger (2014) suggested, the task environment and task initiator are essential factors that modulate writing processes. The task initiator establishes the task environment, specifies the topic, audience, etc., and directs the planner to set writing goals at the control level. Hence, it can be plausible to suggest that task directions, which can be designed to direct someone's attention to another person's

perspective (Apperly, 2018; Apperly, Simpson, Riggs, & Chiavarino, 2006), can be arranged accordingly in order to help writers take audience perspective, increase their audience awareness (Cho & Choi, 2018) and create high-quality reader-friendly (Block & Strachan, 2019; Cohen & Riel, 1989; Kim, 2020; Kim & Graham, 2021) and persuasive passages (Midgette, Haria, & MacArthur, 2008).

Following this suggestion, I will design an audience-oriented experimental writing task that will include pictures with representative faces, vignettes with representative names, a short state PT activity, and a prompt which directly refers to the audience. On the other hand, the non-audience-oriented control task will include pictures without representative faces, vignettes without representative names, a kind of simple brainstorming activity which will not encourage taking reader PT, and a prompt which does not directly refer to the audience (see Methodology section for the details). The experimental task will probably encourage the writers to take audience perspective and keep them in their minds while writing which may require the allocation of EF resources (see Brown-Schmidt, 2009; Schneider, Lam, Bayliss, & Dux, 2012; Wardlow, 2013).

Besides the availability of EF resources, because those who internalized PT skills may show better PT performance (Colman et al., 2017; Woodbridge, 2017) in different contexts (Calvard, 2010; Crocetti et al., 2016; Holt, Bobocel, & Chen, 2021), trait PT (the general tendency of a writer to take others' perspectives) may also influence the efficiency of the experimental task.

Moreover, although this audience-oriented task, which is a kind of state PT activity, may increase the level of engagement with the audience and recruit some of the

limited EF resources (e.g., Davis, Conklin, Smith, & Luce, 1996; Ryskin et al., 2015; Wardlow, 2013), because it may make others' perspectives cognitively more accessible (Ku et al., 2015), having a natural tendency for PT (trait PT) may decrease this dependency of state PT (audience processing) on EF. In other words, high trait PT may have the potential to compensate for the limited EF resources. If they have a natural tendency to take others' perspectives, even the ones with relatively lower EF capability may benefit from this audience-oriented task.

4.2 The aim and contribution

The main aim of this study is to contribute to the literature by designing a task that directs the attention of upper-intermediate level novice Turkish EFL writers to their audience and helps them manage their limited EF resources. The influence of the task, which may be observed in the overall writing quality and the level of persuasiveness of a passage, may be more visible for the participants with relatively higher EF capability and trait PT tendency.

The previous studies indicated the importance of EF (e.g., WM) (Vasylets & Marin, 2021) and PT (Cho et al., 2021), which is associated with audience awareness (Kim & Park, 2019). Again, some studies indicated the importance of embedding the audience component into a writing task (e.g., Block & Strachan, 2019; Cho & Choi, 2018). However, they did not create a writing task that directs writers' attention to encourage a novice EFL writer to consider/maintain her audience throughout writing and investigate its effect on writing performance by considering the EF capability and trait PT tendency. This study may fill this gap and contribute by answering whether a

designed audience-oriented task interacts with EF and trait PT and increases the writing performance of upper-intermediate level novice EFL writers.

4.3 The research questions and hypotheses

This section will report the study's three main questions and the related hypotheses.

Question 1: What are the effects of task type (audience-oriented experimental versus non-audience-oriented control²⁸) and EF capability (high versus low) on the overall writing quality and the level of persuasiveness of the passages written by novice EFL writers?

Hypothesis 1: The experimental audience-oriented writing task will increase the overall writing quality performance of novice EFL writers with high EF capability.

Hypothesis 2: The experimental audience-oriented writing task will increase the level of persuasiveness performance of novice EFL writers with high EF capability.

Question 2: Does trait PT (high versus low) moderate the effect of the task type (audience-oriented experimental versus non-audience-oriented control) on the overall writing quality and the level of persuasiveness of the passages written by novice EFL writers?

Hypothesis 3: The experimental audience-oriented writing task will increase the overall writing quality performance of novice EFL writers with high trait PT tendency.

Hypothesis 4: The experimental audience-oriented writing task will increase the level of persuasiveness performance of novice EFL writers with high trait PT tendency.

²⁸ During the rest of the study, the terms experimental task, experimental condition, or experimental group can be used to refer to audience-oriented task condition, and control task, control condition, or control group can be used to refer to non-audience-oriented task condition.

Question 3: If trait PT moderates the effect of experimental audience-oriented writing task on the overall writing quality and the level of persuasiveness of the passages, can it compensate for the lack of sufficient EF resources when a novice EFL writer with low EF capability was exposed to an audience-oriented experimental task?

Hypothesis 5: High trait PT tendency can compensate for insufficient EF resources (low EF capability) and increase the overall writing quality performance of novice EFL writers exposed to the experimental audience-oriented writing task.

Hypothesis 6: High trait PT tendency can compensate for insufficient EF resources (low EF capability) and increase the level of persuasiveness performance of novice EFL writers exposed to the experimental audience-oriented writing task.

CHAPTER 5

METHODOLOGY

This study was designed as a between-subject experiment that included two critical writing task conditions, an audience-oriented experimental task versus a non-audience-oriented control task. Besides the main experimental writing tasks, the other independent variables were the EF (low versus high) and trait PT (low versus high) capability. The overall writing quality and the level of persuasiveness of the passages were dependent variables.

This section will start with the participant characteristics, the context, the materials, and the procedure of the study. Then, I will focus on the target writing and control measurements. Finally, some ethical considerations will be reported.

5.1 Participants

Participation in the study was voluntary, and this was emphasized on the consent form, which was approved by the Ethical Committee of Boğaziçi University and the private university where the data was collected (Appendix A).

The participants were 133 native Turkish students who had completed the Upper-Intermediate level and just started taking Advanced level English lessons in a private university preparatory school. The mean age of acquisition (AoA) of the sample was 10.96 ($SD = 3.94$), and the mean current age was 20.02 ($SD = 1.056$). The reason for choosing these novice EFL writers were directly related to one of the basic assumptions of this study. Because they are struggling with basic language processes,

their WM (limited EF resources) could be more actively involved compared to relatively more expert writers, and the effect of the experimental task could be more observable (McCutchen, 2000).

Half of these students ($n = 66$) were randomly assigned to the audience-oriented experimental and the rest ($n = 67$) to the non-audience-oriented control group. The gender and department distribution across the groups were similar (See Table 2 for details), and a between-subject t-test analysis revealed no statistically different AoA and current age differences (all P s > .1).

Table 2. The Distribution of Participant Gender and Department Types across Groups

		Experimental	Control
Gender	Female	46	44
	Male	20	23
Department Type	Mathematics / Physical Sciences	20	25
	Verbal / Social Sciences	7	5
	Equal Weight	39	37

Note: In Turkish education system, the students need to take a university entrance exam. In this exam, the weight of the scoring can change depending on the department a student wants to prefer. For example, a student who wants to prefer industrial design or biomedical engineering needs to focus on mathematical (quantitative) part of the exam. However, a student who wants to prefer new media or public relations needs to focus on verbal (qualitative) part of the exam. Finally, a student who wants to prefer psychology or interior architecture needs to focus on equal weight (both qualitative and quantitative) part of the exam.

5.2 Context

5.2.1 EFL (Preparatory School Program) Context

As mentioned above, the data were collected in a preparatory school of a private university. Regarding the school's context, it offered five levels of content-based English learning courses. They were Track 1 (A1), Track 2 (A2), Track 3 (B1), Track 4

(B2), and Track 5 (C1) levels²⁹. Each level lasted one semester (16 weeks), but a student could skip a level if his/her performance was good enough.

In each semester, the students took two language exams (17% of the total), one midterm exam (30% of the total) exam, and one final exam (40% of the total). The language exam included vocabulary and grammar sections. The midterm and the final exam also included vocabulary and grammar, but the main focus was listening, reading, and writing. Additionally, a speaking exam was part of the final exam but was taken the day after. Also, the students were evaluated for their independent studies (5% of the total), for example, for writing their reflections/opinions about a passage processed in the lessons and contributing in this way. Lastly, they needed to complete a portfolio comprising six writing and one speaking task (8% of the total). The writing tasks required the students to use information from the reading or listening passages processed in the lessons (Table 3).

²⁹ According to Common European Framework of Reference (CEFR; Council of Europe, 2001), six main proficiency levels progress from very basic beginner (A1: Breakthrough and A2: Beginner) to intermediate (B1: Lower Intermediate and B2: Higher Intermediate), and advanced levels (C1: Advanced and C2: Very Advanced). In the following part, I will use this framework to give information about the levels.

Table 3. The Distribution of Assessment Components

The Assessed Work	(%)	The Content
Language Exams	17%	Grammar and Vocabulary
Midterm Exam	22%	Grammar, Vocabulary, Listening, and Reading
Midterm Writing Exam	8%	Text-Based Writing
Final Exam	31%	Grammar, Vocabulary, Listening, and Reading
Final Writing Exam	9%	Text-Based Writing
Final Speaking Exam	8%	Sharing ideas in pairs or groups, etc.
Independent Studies	5%	Contribution to the lessons, etc.
Portfolio	8%	6 writing and 1 speaking tasks

Completing at least four tasks in this portfolio (6 writing and one speaking tasks) and attending 85% of the lessons each semester were the basic requirements to pass to the following levels successfully. Otherwise, they were asked to repeat the same level in the following semester. If the students from the first four levels completed these requirements and got an overall average between 0-64, they would move to the next level. For example, a B1 student who got 64 would move to B2. However, if a student from the first three levels completed these requirements and got an overall average between 65-100, she would move to two levels above. For example, a B1 student who got 65 would move to C1 (see Table 4 for a sample scenario).

Table 4. A Sample Possibility Table for a B1 Level Student

The Current Level	Conditions	The Next Level
B1	Less than 4 Portfolio Task Less than 85% Attendance	B1
	0-64	B2
	65-100	C1

When it comes to the successful completion of the preparatory school, there were two alternative ways. The first was to directly pass the school by taking sufficiently high scores from midterm and final exams in B2 or C1. The average of these exams had to be 75 or above for B2 and 65 or above for C1 students. The second alternative was to gain a right to take the proficiency exam of the school, which consists of reading (40%), listening (30%), and writing (30%) parts, and to be successful in it by getting 60 or over. To have this opportunity, the student's overall average had to be 75 or above for B1, 65 or above for B2, and 60 or above for C1 students (Table 5).

Table 5. The Conditions to Complete the Preparatory School Successfully

The Current Level	Conditions	Proficiency Exam/ Direct Pass	Proficiency Score Needed
B1	The overall average is 75 or above	Proficiency Exam	60 or above
B2	The average of midterm and final is 75 or above	Direct Pass	---
B2	The overall average is 65 or above	Proficiency Exam	60 or above
C1	The average of midterm and final is 65 or above	Direct Pass	---
C1	The overall average is 60 or above	Proficiency Exam	60 or above

The students who took English lessons during two semesters (32 weeks) after enrolling at the university and met the criteria but failed to pass their departments were seated in a B2 or C1 classroom in the summer school. The target sample of this study was those C1 students³⁰ who were required to attend the summer school program for

³⁰ Their attendance level was similar across groups.

eight weeks; the data collection period corresponded to the first six weeks of this period.

The initial starting level of these students (at the beginning of the academic year) was mostly A2 ($n = 90$). However, there were students whose initial seated level was A1 ($n = 31$), B1 ($n = 3$), and B2 ($n = 7$). Importantly, the distribution of the participants across the conditions was similar to each other (Table 6).

Table 6. The Distribution of the Initial Levels across Groups³¹

	Experimental	Control
A1 (Track 1/Elementary)	17	14
A2 (Track 2/Pre-Intermediate)	44	46
B1 (Track 3/Intermediate)	2	3
B2 (Track 4/Upper-Intermediate)	3	4

5.2.2 Writing Context

To begin with, the school did not adopt a traditional approach that emphasizes the structure of paragraphs or essays. For example, the students were never told to create a thesis statement in the introductory paragraph, etc. However, there was an emphasis on task fulfillment, language use, and the rhetorical pattern, which can be associated with different paragraph/essay types (e.g., description, cause-effect, compare-contrast, advantage-disadvantage, or problem-solution).

To illustrate, the main course book for A2 students was Unlock 3 Reading, Writing and Critical Thinking (Westbrook, Baker, & Sowton, 2019). In the program, the

³¹ I did not prefer excluding those students who started at higher level in order not to lose data. However, I will check the influence of the initial level of the participants if needed.

teachers were asked to omit the writing part of the second unit, which was about identifying the main, supporting, and concluding sentences in a short paragraph. Again, a part that focused on identifying a topic sentence in the third unit, writing topic sentences in the fourth unit, writing concluding sentences in the seventh unit, and essay organization in the eighth unit were asked to be omitted during the lessons. However, the writing task of the arguments for and against funding space exploration, completing a T-chart with arguments for and against, evaluating the arguments to be used, and some related phrases were included.

Beyond this general approach, the writing context can be thought of in four categories. The first one was six writing portfolios. These tasks asked the students to consider a listening or reading passage from the lessons and to use the information in it, rather than personal ideas, to answer a writing question. For example, the fifth portfolio task of A2 level was “Write ONE effect and TWO solutions for traffic congestion using the information from the UNLOCK 3 RW Unit 3, Reading 2” ... “Key words and linkers you may use are: practical, lead to, it is important to ..., etc.”.

The second category can be the midterm and final writing exams. They followed a similar procedure, and the answers were expected to base on the processed passages rather than personal opinions. There were three important criteria in the evaluation of these passages. They were “Task Fulfillment” (whether the question was answered with details), “Use of Language” (whether appropriate vocabulary and sentence structures were used), and “Rhetorical Pattern” (whether the ideas follow each other smoothly and the linkers were used accordingly). The rubrics were not contextualized for the tasks, and these general criteria were used to evaluate all the passages written to answer a

portfolio or an exam question. Apart from these, the students were taught and were strongly encouraged to paraphrase the sentences whenever they used lesson material as a resource text. Starting from the B1 level, their final writing exam texts were uploaded to Turnitin, and if the similarity rate was 30 or above, they got 0.

The writing tasks the teachers created can be considered another category. These tasks generally aimed to collect the personal opinions of the students about the materials processed during the lessons, and their structure was generally dependent on the needs and expectations of the students. For example, depending on the level of the students, I tended to ask students to write a short paragraph about the issue of that day's lesson before coming to the classroom. Again, when I asked my colleagues whether they ever asked the students to write short or long opinion/argumentation passages, one of them reported that "They are always short. I ask questions like Do you agree with the lecturer/author/the person in the video/ this particular woman or man? and give at least two reasons (these HAVE TO come from credible sources) why you agree or disagree with the ... or What would you do if you were...? Would you do ... if you were ... Explain why fiving reasons (again HAVE TO be credible)". The other one said that "I ask the students, Do you agree/disagree with the main idea given in the text?/Do you agree/disagree with the writer of the text? Please write your response agreeing or opposing the necessary details from the text.". These kinds of activities depended on the teachers' approach, and there were no lessons specifically allocated for teaching writing. It was incorporated into teaching the other skills.

The final category is the proficiency exam and its writing tasks which many students need to take in order to pass the preparatory school. The writing part of this

exam included questions such as “In your opinion, what are the most important factors that make a new product successful? Write an essay of at least 350 words. Give reasons and examples.” (Kadir Has Üniversitesi, n.d.). The students were given some possible ideas, such as “customer needs,” “prices”, etc., and were told they could use them. They had 10 minutes to complete the compulsory brainstorming and outlining part and 50 minutes to write the essay. The criteria for evaluating the essays were similar to the portfolio and the other exams, which included task fulfillment, rhetorical pattern, and language use.

5.3 Materials

As detailed in the procedure section below, I separated the study into three stages. The background data, which included personal information and diagnostic writing, was collected in the pre-experimental stage. Then, within two weeks, the experimental essays and the retrospective questionnaire data were collected. After that, within around 3-4 weeks, EF data (AC and WM), and finally, a questionnaire which investigated strategy use during EF tasks was applied and an online trait PT scale data was collected (Table 7).

Table 7. The Material List

Stage	Material	Aim
Pre-experiment	Background Questionnaire	Collecting general background information
	Diagnostic Writing	Comparing EFL writing level of the participants
Main-experiment	Main Experimental Writing	To assess the effect of directing the attention of the writer to the audience
	Retrospective Questionnaire	To collect reflections on the main writing and get some background data
Post-experiment	AC and WM	To evaluate the EF capacity of the participants
	Trait PT Scale	To evaluate the PT tendency of the participants
	EF Tasks Strategy Use	To investigate what kinds of strategies used by the participants

*In addition to these the language related data (initial proficiency, reading, vocabulary, and grammar scores) were collected from the testing unit of the school.

5.3.1 Pre-experiment stage materials

5.3.1.1 Background questionnaire

A background questionnaire that aimed to collect basic information, such as age and student department, was employed before the main experiment (Appendix B).

5.3.1.2 Grammar and vocabulary exam of the preparatory school

Although the participants of this study were from the same level, there could be some variations in their language background. To control this in statistical models, I collected the summer school vocabulary and grammar exam scores of the students from the testing unit of the school.

5.3.1.3 The initial language proficiency exam of the school

Again, to check the language background of the participants, I used the initial language proficiency exam results of the students taken at the beginning of the academic year. The test consisted of four sections that employed increasingly difficult language use questions and aimed to detect the students' proficiency and seat them accordingly.

5.3.1.4 Reading comprehension exam of the school

Because the students needed to read some vignettes during the main writing stage (see below), and reading ability can be crucial in writing (e.g., Kim & Graham, 2021; Kim & Park, 2019) for different reasons (e.g., reading the written text so far to evaluate it, reading a task, etc.) (Hayes, 1996), checking their reading comprehension performance would be important. For this purpose, I used the detailed reading scores that the students obtained in the final exam of the summer school.

5.3.1.5 Diagnostic writing task

As Rahimi and Zhang (2018) did, I wanted to consider/control the writing performance of the participants, as well. For this reason, I employed a diagnostic writing task (Appendix C). The aim was to check whether the students' writing performance was similar across the conditions. Moreover, because the organization of the main question part of this task was similar to the main writing, it would familiarize the students with the main writing task.

5.3.2 Main-experiment stage materials

5.3.2.1 Main writing tasks

The main experimental data was collected via the prompts, which instructed the participants to write a convincing opinion passage by giving reasons and details (Appendix D). This type of passages can be associated with argumentative or persuasive essays, which require the writer to be highly aware of the audience and aim to convince others to take some action etc. (Ferretti, MacArthur, & Dowdy, 2000; Schneer, 2014). They employ many elements highly associated with thinking processes and, compared to narratives, require relatively more cognitive demand (Kim & Pae, 2021).

The writing tasks, except from the writing itself, had four components: vignettes, pictures, a kind of simple brainstorming (non-audience-oriented control) or PT (audience-oriented experimental) section, and a prompt. Firstly, they included a short message/vignette, defined as carefully created realistic scenarios to enhance experimental realism, thus improving internal and external validity (Aguinis & Bradley, 2014). The vignettes are considered efficient and less disruptive data collection tools (Wilson & While, 1998) which may be used, for example, for encouraging PT and increasing the originality of ideas during brain storming (So & Joo, 2017). When it comes to its format, different video-recorded vignettes (Mohr, Howells, Gerace, Day, & Wharton, 2007) or written vignettes (Gliner, Haber, & Weise, 1999; Hanson & Scott, 1995) can be presented, and the reactions to these passages can be collected. Important for our present concern, they are successfully used in PT studies (e.g., Holt et al., 2021; Mohr et al., 2007). For example, a video-recorded vignette was used, and the following prompt manipulated PT (e.g., self-versus-other) before asking for a (written) response

(Davis et al., 2004), different written vignettes were presented to assess the reader's awareness of the victim in the passage (PT reactions) (Hanson & Scott, 1995), and so on. Essentially, they can manipulate/control independent variables and assess dependent variables (Aguinis & Bradley, 2014). Thus, presenting vignettes can be appropriate to direct the attention of the writers to the audience. In this study, I employed written (to be read) vignettes as a part of the main writing task.

To prepare and finalize these vignettes, I searched the internet for the issues and activities related to the student clubs (theatre, music, and outdoor sports) and wrote the first drafts. Like studies that manipulated the independent variables via vignettes (e.g., Skarlicki & Turner, 2014), I tried to keep the core content constant across the conditions. The vignettes were around 80-90 words. In the vignette with audience focus (experimental condition), there was an emphasis on people, the sentences were active, and the names of the participants or the words that directly refer to them were used. For example, the names of the representatives were explicitly stated, and the activities, expectations, and needs of the club members were mentioned in active sentences. Each vignette started with a general introduction. Then, the aim and interests of the club members were stated. Some activities and reasons were mentioned, and a need was expressed. The emphasis was on a club rather than people in the non-audience-oriented (control condition). Instead of the names of the participants or the words that refer to them, the words such as 'club' were used. The topics were again related to the three student clubs mentioned above.

After creating the first drafts, I asked an expert (my thesis advisor) to get feedback about both the internal validity (e.g., whether there is a consistency between

the research aim, methodology, and the written vignettes) (Skilling & Stylianides, 2020) and external validity (e.g., whether the content is realistic, etc.) (Bradbury-Jones & Herber, 2014; Skilling & Stylianides, 2020). Then, I showed the passages to the instructors³² who taught the target group of EFL learners and got their feedback about the appropriateness of the language and their ideas about whether my manipulation can successfully induce audience focus, etc. Finally, I also got feedback from the target-like students by asking them to evaluate the following statements. ‘The content of this passage was plausible’, ‘The complexity level of this passage was normal, it was not difficult to understand’, ‘The passage sounded natural enough’, ‘I am familiar with this audience (student clubs)’, etc. They assured that the vignettes related to the school clubs³³ were appropriate/understandable, etc.

Secondly, I presented pictures next to the vignettes because adding visuals can strengthen engagement (Lee & Goh, 2020). The visuals included three to four human faces in the vignettes with audience focus, but general topic representative pictures in the vignettes without audience focus. All the pictures were cropped in a one-to-one ratio (Appendix E).

Thirdly, I embedded a state PT (experimental) and a kind of simple brainstorming (control) activity into the task. The PT literature indicated that PT increases self-other overlap in cognitive representations (Galinsky et al., 2005) and the other's perspective (the audience in this case) was processed more when it was relevant

³² I thank to my colleagues, Bassel Kaskas and Gülsen Yılmaz.

³³ There was another alternative topic, city projects. I excluded it because the students reported and the pilot study results revealed that student clubs was a more appropriate topic than that; they found the city-projects topic boring, etc.

to the task instructions (Apperly et al., 2006). Moreover, it has some benefits such as going beyond the mental routines and involving more cognitively demanding activities (Ku et al., 2015), which may result in, for example, remembering more details (e.g., after shifting the perspective) (Anderson & Pichert, 1978). Accordingly, this part aimed to differentially increase the students' engagement with the audience's perspective in the experimental condition. For similar purposes, some researchers asked the participants to write down about, for example, a typical day by taking the perspective of another (e.g., a photographed) person or seeing the world from her eyes in the experimental condition, but just to write down about the typical day of the person in the control condition (Galinsky et al., 2005). Inspired by this, I created two versions of this part. The following instructions were presented under the vignette and picture: “Now, please take the perspective of these club members (their expectations, thoughts, feelings, etc.) and write whatever comes to your mind. You have 3 minutes.” On the other hand, there was no explicit direction of attention to the audience in the control condition, and the instruction included, “Now, please just write whatever comes to your mind. You have 3 minutes.”

The last component of the tasks was the essay prompts adapted from Rahimi and Zhang (2018). The prompts were previously used in the audience awareness literature to emphasize and manipulate the audience focus. For example, Cho and Choi (2018) emphasized the audience (“Your report is intended for your classmates, who have not read or heard the sources, may not know anything about the topic and are depending on you to be clear and informative”) or the text itself in the prompts³⁴. Departing from these

³⁴ Taken from the supplementary material of Cho and Choi (2018) for the sake of illustration.

kinds of samples and thinking of it can increase the efficiency of the manipulation further, I also organized the prompts. With this purpose, they employed either an audience-oriented task (experimental) or a non-audience-oriented task (control) content by referring to people (rector and club members) or to organizations (management and clubs), respectively (Appendix D).

After piloting on myself, I also piloted the whole experiment on the target-like students to check the appropriateness of the tasks. The aim was to understand whether the instructions were clear, whether the manipulation seemed to work, and whether there were any procedural problems, such as the duration needed to complete the task. Additionally, I asked two target-like students to translate the tasks into Turkish to check their comprehension.

The writing task I reported above was the final version I created after the first pilot and translation stage. After piloting the experiment with 12 students and asking two students to translate them (to check their comprehension), I got their ideas about the whole writing task, which at that time included an alternative topic (city projects). After these checks, I gave up the alternative topic and a control task with no brainstorming part. Both proved that they would make the design unnecessarily more complex.

Around one month after the first pilot, I conducted a second pilot with two students. My observations revealed that the tasks and the other materials were ready for the experiment with respect to time management, etc.

5.3.2.2 Retrospective questionnaire

A brief retrospective questionnaire, which is less intrusive and more practical than think-aloud and introspection (Ong, 2014), was used to check the interaction of the participants with the task (e.g., to what extent they considered the audience in their writings, etc.) and control the similarity of the groups with respect to their writing and topic background (Appendix F).

This retrospective questionnaire included a manipulation check inspired by Todd, Bodenhausen, Richeson, and Galinsky (2011). The participants were asked to mark whether they tried to take the perspective of the club members (e.g., when I read their messages, I tried to understand the expectations of outdoor sports club members from their point of view.) and whether it was difficult to take their perspectives (e.g., when I read their messages, I had a hard time understanding the expectations of outdoor sports club members from their point of view.) on a five-point scale. They did this for each club and then reported whose perspective they took. An adapted item from Smith (1995) “Did you think about a reader or readers while writing your essay? Who? Why did you choose this reader or these readers?” (p.418) was used for this purpose. This part was followed by a part that included questions related to the main writing (difficulty in understanding the prompt, etc.) and a part that tried to understand the approach of the participants to the evaluation of their writings (How do you think this text you wrote will be evaluated? Can you give details?). Then, their interests in the clubs were checked by asking questions such as: Do you have a special interest in music, theater, or outdoor sports? Can you give details? These sections were followed by sections that questioned the participants’ previous writing experience, general writing experience, and free time

writing activities (see Appendix F for the complete questionnaire and the Control Measurements section below for coding details).

5.3.3 Post-experiment materials

5.3.3.1 Attention control and working memory tasks

5.3.3.1.1 Flanker (AC) task

The idea behind the flanker task is based on Eriksen and Eriksen (1974), who indicated that incompatible letters around a target letter create noise, distract attention, and increase reaction time.

In this task (Martin et al., 2021), a fixation point is presented at the beginning of a trial. Then, a target arrow that points either to the left or right and is accompanied by two flanking arrows on both sides is presented at the center of the screen. There are two different trial types. In a congruent trial, the flanking arrows point in the same direction as the target (e.g., $\leftarrow \leftarrow \leftarrow \leftarrow \leftarrow$). On the other hand, in an incongruent trial, they point in the opposite direction (e.g., $\leftarrow \leftarrow \rightarrow \leftarrow \leftarrow$). The task is to press a button (“z” for the left and “.” for the right) as fast as possible to show the direction of the target/middle arrow. The participants keep their index fingers on the keyboard buttons to respond quickly and accurately.

There are 144 (96 congruent and 48 incongruent) trials with a randomized 400-700 msec inter-trial interval. The dependent variable, the flanker interference effect, is calculated by subtracting the subject’s mean reaction time in congruent trials from that in incongruent trials. Only accurate trials were considered in this calculation.

Although this is a task widely used to assess attention, Draheim, Tsukahara, Martin, Mashburn, and Engle (2021) reported that using reaction time to assess AC can be problematic because it is sensitive to speed-accuracy interactions. Moreover, some participants may be more deliberate and slower respondents, so their performance may seem to be worse than those who are quick but error prone. Thus, we cannot be sure whether it is about general processing speed or AC. Because of this, I included an alternative AC (antisaccade) task, which I will explain in the next part.

5.3.3.1.2 Anti-Saccade (AC) Task

Antisaccade task is a nonverbal attention task that requires minimum memory retrieval (Kane et al., 2001). In each trial of it (Draheim et al., 2021), a fixation cross is presented for between 2000-3000 msec and this is accompanied by an alerting tone for the last 300 msec. After that, an asterisk cue appears on the 12.3° visual angle left or right of the central fixation for 300 msec. Following this, a target letter, “Q” or “O”, is presented on the opposite side of the asterisk for 100 msec and is immediately masked by “##” for 500 msec. The task is to look at the opposite side of the asterisk in order to catch the target letter (“O” or “Q”) and to press the associated letter on the keyboard. Following the response, correct/incorrect feedback is presented for 500 msec. There is a 1000 msec blank interval between the trials (Fig. 8). The participants complete 72 target trials in total. The dependent variable is the level of accuracy in reporting the presented letters (Draheim et al., 2021).

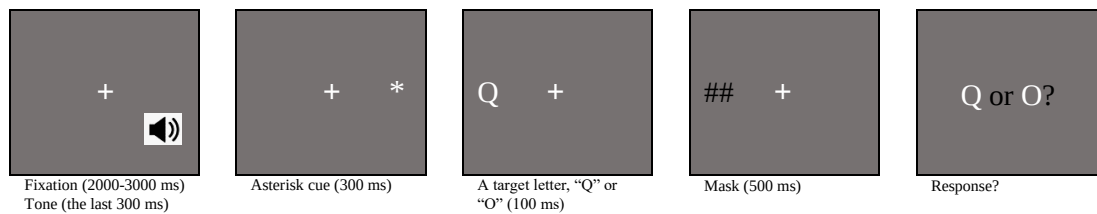


Fig. 8 A sample anti-saccade task trial

To perform in this task well, the participants must inhibit a dominant response (e.g., looking at the asterisk) and plan and execute a voluntary saccade in the opposite direction. Because these anti-saccades require not only inhibition but also the direction of the focus and are controlled voluntarily (top-down), they require the inclusion of executive control processes, which are associated with WM capacity (Unsworth, Schrock, & Engle, 2004).

As mentioned above, the reason for including this task is that it can be less prone to speed-accuracy trade-offs. In this task, the participants need to make a quick saccadic movement to the opposite direction of the cue (asterisk) in order to identify a target stimulus (Q or O) that is presented very shortly (100 msec) and masked quickly. If they are not quick enough, they will miss the target letter and will not be able answer which one is presented. Thus, the speed is irrelevant, and, unlike the flanker task, there is a minimum speed-accuracy trade-off. This makes this accuracy-based anti-saccade task a valuable alternative for assessing the AC of the participants (Draheim et al., 2021).

5.3.3.1.3 Symmetry Span (WM) Task

In symmetry span task, which assesses WM (Unsworth, Redick, Heitz, Broadway, & Engle, 2009), the participants are required to keep the location of a series of squares presented in a 4x4 matrix as they decide whether a presented pattern is symmetrical or

not (Fig. 9) (Draheim et al., 2021). In the task, first, a 16x16 matrix, which includes white and black squares, is presented, and the participants decide whether the pattern is symmetric. To be sure that they make these symmetry decisions accordingly, the participants are asked to have at least 85% accuracy and to follow their scores in the upper-left corner of the screen. The aim is to keep them engaged and prevent rehearsing of the target boxes. Then, a 4 x 4 matrix is presented, and one of the to-be-recalled squares is highlighted in red. The number of alternations on each trial depends on the number of spatial locations to be recalled (3 to 7). Finally, at the end of a trial, the participants must click on the locations in the previously presented order. The dependent variable of this task is the total number of locations recalled in the correct order (the partial span scores) (Conway et al., 2005).

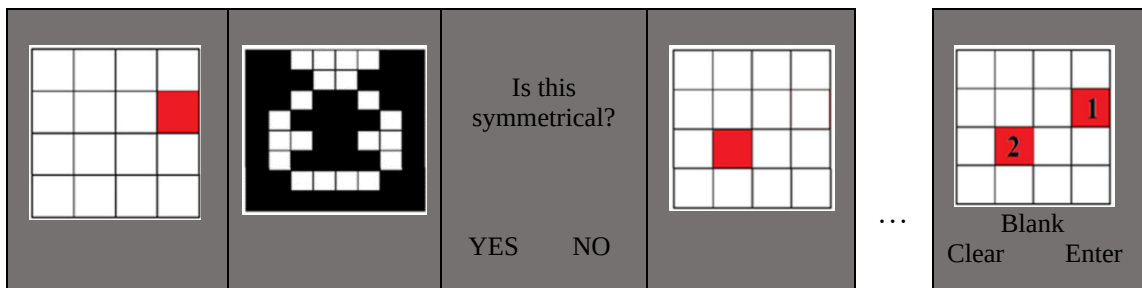


Fig. 9 A sample symmetry span task trial

5.3.3.1.4 Operational span task

I also used a shortened version of the Operational Span Task (Foster et al., 2015; Unsworth, Heitz, Schrock, & Engle, 2005) to assess WM capacity. This task requires participants to keep a series of letters in their memory as they solve simple mathematical problems (Fig. 10) (Draheim et al., 2021). In the task, first, a math problem (e.g., $(2 \times 1) + 1$) is presented as a distractor, and the participants need to decide whether the

presented result in the following window is true. Similar to symmetry span task, they also need to reach at least 85 % accuracy. Following that, a to-be-recalled letter (e.g., “F”) is presented. Again, the number of this alternation on each trial is dependent on the number of letters to be recalled (3 to 7). Finally, at the end of each trial circle, the participants need to select the letters in their presented order. The dependent variable is the sum of the number of letters recalled in the correct order (the partial span scores) (Conway et al., 2005).

F	$(1*2)/2=?$	3	P
		TRUE FALSE	

...

1	F	H	J
	K	L	N
2	P	Q	R
	S	T	Y
Blank			
Clear		Enter	

Fig. 10 A sample operational span task trial

5.3.3.1.5 Retrospective Strategy Use Self-Reports

After collecting the EF data, the participants were asked to report if they used any strategies during the tasks. The answers to the questions such as "did you use any strategies as you completed the attention task during which you chose the direction of the arrows (left or right)?" were collected via an online form. The aim was to understand whether they did something unnatural and whether they could influence the obtained results. For more details, please see the “Control Measurements” section below.

5.3.3.2 Trait Perspective-Taking Scale

Finally, online trait PT scale data were collected to get insight into the participants' general PT characteristics and to see whether they are associated with their writing

performance. For this purpose, I used the Turkish version (see Engeler, 2005; Bolat, 2019) of Davis' (1980) 7-item, 5-point PT scale, which includes items such as “I try to look at everybody's side of a disagreement before I make a decision” (Davis, 1980, p. 7; see Appendix G for the other items).

5.4 Procedure

The participants of this study were recruited on a voluntary basis. Before the experiment, their permission was taken via a consent form, and following the experiment, they were paid 50 liras (~3 dollars) for their participation.

The pre-and main-experiment stages were paper-based, and the data were collected in the classrooms during the lessons. The AC and WM data were collected in a laboratory with four computers in separate sections. Lastly, trait PT data were collected by using online Google Forms.

The pre-experiment session lasted around 75 minutes. After summarizing the whole procedure and the overall aim of the study, the consent forms were given, and three-digit participant codes were assigned to the subjects who signed the form. From this point on, only the participant codes were associated with the data.

Following the collection of background questionnaires (Appendix B), the participants had 40 minutes to write an essay in response to the diagnostic writing task (Appendix C).

The instructions and the procedure of the diagnostic writing part were as the following. First, the participants were reminded that they mustn't use their phones during writing. Then, they were given 3 minutes to read the writing task instructions.

Following that, they were encouraged to ask clarification questions if they had any. To be sure that everybody understood the task, I also assessed their understanding by asking them to evaluate the statement “I have understood the instructions of the writing task.” on a 5-point scale (1 meaning *I have not understood at all* and 5 meaning *I have completely understood*.). If any of them evaluated it by giving 4 or a lower score, the task was clarified again. The most important clarification at this stage was about the employment of their own ideas. Because the main aim was to collect their opinions via an English text, I emphasized that they would share their ideas/opinions about the issue³⁵ by writing around the 250-word long text. After being sure that the task was understood, they were reminded that they could use the empty part of the writing task paper to do any preparation (planning, creating an outline, etc.). They were also told they needed to use a pen but could cross out any words, sentences, or paragraphs and revise their writing anytime. Then, two empty papers were distributed to collect the essays from the participants. They were given 40 minutes, and the timing was reminded from time to time (the last 20, 10, and 5 minutes). The writing duration of each participant was noted down.

In the main experiment stage, at least two days after the pre-experiment session (within two weeks), the main writing essays and the following retrospective questionnaire data were collected. Similar to the diagnostic writing, this paper-based data was collected from the classrooms randomly assigned as experimental or control groups.

³⁵ Apart from this, similar to, for example Qin and Liu (2021), I did not encourage the participants to avoid using the information given in the vignettes.

The main data collection started with the writing task (Appendix D). Although the general procedure of this task was similar to the diagnostic writing, there were some differences, some of which were specific to control or experimental groups. First, all the participants were asked to read the instructions carefully for 2 minutes. Then, the three vignettes/messages were counterbalanced and distributed one by one, and the participants were told that they had 1 minute to read the passages. After reading each vignette, they were asked to reflect on the vignettes in a kind of brainstorming activity within 3 minutes. However, there was a difference in the instructions given to the groups at this time. The control group was just asked to note down whatever came to their minds (relevant/irrelevant). On the other hand, the experimental group was encouraged to take the perspective (which was simply defined as trying to understand the expectations, needs, and feelings of other people) of the audience (rectorate and student clubs) and note down whatever came to their minds. The completion of this part took around 15 minutes.

Following this, the writing prompts were distributed, and the participants were given 2 minutes to read them by themselves. Similar to the diagnostic writing, their comprehension was checked. Again, it was emphasized that the main aim of writing was sharing their opinion about the issue (the allocation of money) by writing an English text. The empty writing papers were distributed after ensuring they understood the writing prompt. The whole task processing (including task introduction, vignette

reading, a kind of brainstorming, and prompt processing) lasted around 20 minutes. Finally, the participants were given 40 minutes to write their target essays³⁶.

After completing their passages, the participants took ten minutes break and completed the retrospective questionnaire (Appendix F). This main experimental stage lasted around 75 minutes on average.

When it comes to the post-experimental stage, as mentioned above, it mainly consisted of EF (WM and AC), trait PT, and strategy use during EF tasks data collection. Following the main experiment stage, the participants were invited to the laboratory, and their WM and AC data were collected in two separate sessions. The counterbalanced flanker and symmetry span tasks were given in the first, and the counterbalanced antisaccade and foster operation span tasks were given in the second session. Four computers in the laboratory inserted into four separate cubicles were used for this purpose. This allowed me to give initial instructions, at most, to four participants simultaneously and to collect their data separately. They were actively followed to check if they understood the task and tried to do their best to complete it accordingly (all the computers were visible simultaneously). Each session lasted around 30 minutes. Upon completing the data collection, the participants were paid 50 liras (~3 dollars) and thanked once again.

Finally, the online data for strategy use during EF tasks were collected, and in average around one month after the first session, I created a Google form and collected trait PT data using Davis's (1980) 7-item PT scale (Appendix G).

³⁶ Following pilots, I decided that these may be enough for participants to process and complete the writing task. This was important because, as Becker (2006) reported, the time factor can create extra cognitive load and I wanted to avoid its contribution.

The whole data collection was completed within six weeks. You can see the summary of the whole data collection procedure in the flowchart below (Fig. 11).

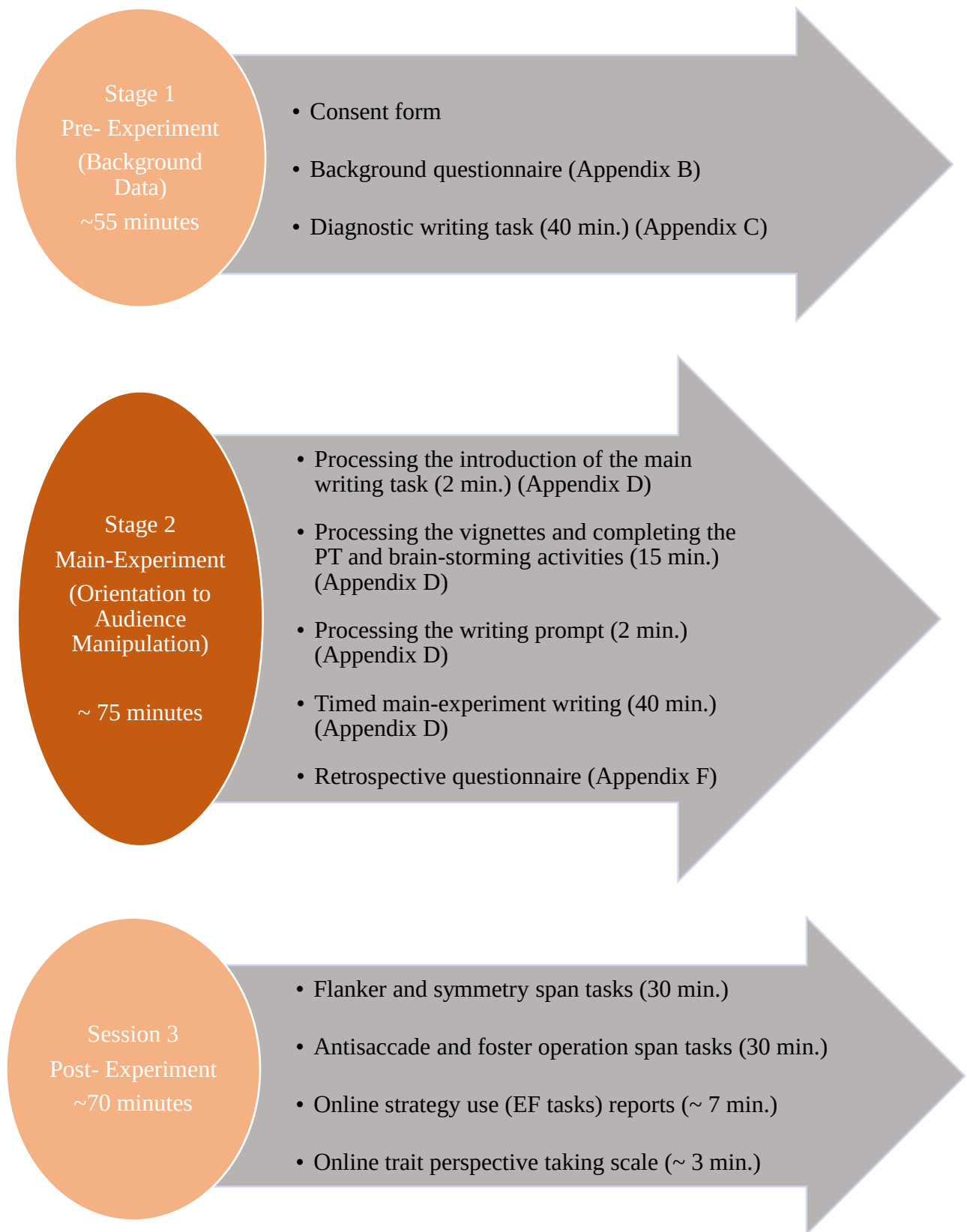


Fig. 11 A summary of the data collection procedure

5.5 Target writing and control measurements

This section aims to report the evaluation and coding procedure of the data collected from the participants but needs further processing (e.g., main writings and retrospective self-reports). I will consider them under three titles. The first one will focus on the main experimental measurements of the study: the overall writing quality and the level of persuasiveness. The second part will focus on control measurements, including the measurements which aimed to check the background similarity of the groups, the reflection of the groups on the writing task, and the strategies they used during the EF assessment. Lastly, the third part will consider the control measurements, which aim to assess the efficiency of the manipulation. These data will be analyzed later, under the preliminary analysis part of the results chapter.

5.5.1 Main experimental writing measurements

5.5.1.1 The overall writing quality

The overall writing quality of the passages was assessed by two independent evaluators³⁷ on the basis of the writing assessment rubric of Jacobs, Zinkgraf, Wormuth, Hartfiel, & Hughey (1981) (Appendix H).

This rubric is widely used in the EFL literature (Weigle, 2002), for example, to evaluate the overall quality of argumentative essays (e.g., Ong & Zhang, 2010; Ong, 2013; Rahimi, 2019; Xu, Zhang, & Gaffney, 2023). Additionally, its subcomponents are used to evaluate argumentative essays (e.g., content and organization) (Rahimi, 2019), and introductory parts of narratives (Rahimi & Zhang, 2021), etc.

³⁷ Special thanks to Mehtap Güven Çoban for her valuable support as a second evaluator.

It is also used to classify the proficiency of EFL writers (Jacobs et al., 1981; Ong, 2014; Ong & Zhang, 2010; Rahimi, 2019). Jacobs et al. (1981) suggested that the scores between 74 and 100 can be classified as advanced (74-82: low advanced, 83-91: advanced, 92-100: high advanced), the scores between 47 and 73 can be classified as intermediate (47-55: low intermediate, 56-64: intermediate, 65-73: high intermediate), and the scores 46 and below as beginning (0-37: beginning, 38-46: high beginning). As an alternative, some researchers considered their study's context and classified the scores accordingly. For example, Ong (2013) classified participants as high-proficiency writers if their scores were between 68 and 75, as mid-proficiency writers if their scores were between 57 and 65, and as low-proficiency writers if their scores were between 46 and 56.

The rubric considers five aspects of writing (content, organization, vocabulary use, language use, and mechanics). Each of these aspects has four levels. The first two levels, excellent to very good and good to average, are associated with successful communication, while the last two levels, fair to poor and very poor, are associated with communication breakdown. Thus, effective communication is at the center of the whole evaluation.

The evaluation procedure was as the following. After studying the operationalizations and grading the sample papers in Jacobs et al.'s (1981) guidebook, the raters discussed the purpose of the writing task. There were some context-dependent things to consider. For example, as mentioned in the context section, the school did not encourage the participants of this study to organize the essays by dividing them into introductory and conclusive paragraphs, etc. Thus, during the evaluation, it was

important to focus on the organization of the ideas, the supporting details, etc., rather than considering the mechanical division of the paragraphs. This would be in accordance with the “being succinct” criteria of the organization, which meant the concise direction of all ideas to the central focus of the essay. Other than that, as Jacob et al. (1981) stated, all the papers were to be evaluated on their own merits rather than in comparison with the other ones. Each evaluation was expected to take 2-3 minutes in a distractor-free environment. Again, after evaluating around 25 essays, the rater had to stop and check the operationalization of the criteria. If a need was felt, she/he was recommended to check the sample essays and sample evaluations and to do a consistency calibration with the other rater.

In the next step, I randomized all the essays by giving them a random code between 401 and 999 and chose some anchoring (seemingly poor, moderate, and strong) papers. Then, the second rater and I evaluated these papers to check if we were calibrated. Seeing that we agreed on the operationalization of the criteria and their employment in our assessments, we continued evaluating the papers separately.

Each paper was evaluated by each rater at least twice. In the first evaluation, the focus was on the content and the organization. In the second cycle, the focus was on evaluating vocabulary knowledge, language, and mechanics. Importantly, all these components were considered to the extent they strengthen or weaken the communication because they were thought of as different windows/viewpoints for judging the overall communication. In other words, the important question was to what extent the writer effectively communicated a clear and complete message. For example, the language

evaluation was not about the number of syntactic errors but whether a syntactic error affects the comprehension of meaning.

After evaluating all papers, an inter-reliability analysis was conducted. The computed Cronbach's alpha was high for both diagnostic ($\alpha = .91$) and main writing ($\alpha = .89$). Still, the scores were compared, and all discrepancies (more than 10 points difference) were resolved by negotiating. The final interrater reliability was increased to .94 for both diagnostic and main writings. These values are above the desirable reader reliability (.85) reported by Jacobs et al. (1981, p. 39). At the final step, the average scores of the evaluators were taken for descriptive and inferential analysis.

5.5.1.2 The level of persuasiveness

For the evaluation of the written passages with respect to their persuasiveness, I adopted a holistic approach and contextualized a rubric considering the previously created ones (e.g., Ferretti et al., 2000, Qin & Karabacak, 2010, Ong, 2013, Vögelin, Jansen, Keller, Machts, & Möller, 2020) (Appendix I).

I reviewed and analyzed several holistic assessment tools shared in the literature in order to adopt some operationalizations and adjust them to create a holistic rubric that can evaluate the general level of persuasiveness. The main reason for this adjustment was that most of the rubrics in the literature (e.g., Ferretti et al., 2000, Ferretti, Lewis, & Andrews-Weckerly, 2009) focus on typical agreement/disagreement argumentative/persuasive essays, which include instructions that directly ask the participants to take a clear stand and write a persuasive essay. However, in our institution, although the students were encouraged to support their ideas by including

reasons and considering different perspectives (e.g., advantages and disadvantages of an issue), etc., they were not mechanically taught to write traditional argumentative (or any other) essays (starting with an introduction, including a thesis, etc.). Considering the research in the literature (e.g., Rahimi, 2019; Revesz, 2011; Watson, DeSanctis, & Scott Poole, 1988) and considering the school's context, I created a task which asked for opinions on how to share a limited amount of money, 80.000 liras among three student clubs (outdoor, theatre and music). The participants needed to decide how to allocate this money and write a convincing passage by providing reasons and examples. Thus, the contextualization for assessing persuasiveness was important to reflect the demands of this untraditional task given in this specific context³⁸ (e.g., Tedick, 2002; Varghese & Abraham, 1998).

Accordingly, I considered the holistic rubrics in the literature (e.g., Chase, 2011; Cheng & Chen, 2009; Ferretti et al., 2000; Geng, 2017; Majidi, Janssen, & de Graaff, 2021; Nussbaum & Kardash, 2005; Nussbaum & Schraw, 2007; Reznitskaya, Kuo, Glina, & Anderson, 2009; Ong, 2013; Qin & Karabacak, 2010; Wang, 2016) and adapted the basic components of them in order to create a 7-level rubric and to evaluate the passages from a specific dimension, persuasiveness. To illustrate, one of the holistic rubrics related to our present concern is the persuasive writing rubric introduced by Ferretti et al. (2000). The levels of it range between 0 (no opinion is stated) and 7 (well elaborated and addresses the opposition). To get the highest scores, the participants need to write a well elaborated persuasive passage that employs clearly explained opinions and supporting/convincing reasons. This highly evaluated passage would also employ

³⁸ Beyond these, another reason for adjusting the rubric was that the persuasiveness rubrics in the literature are mostly used to assess L1 rather than L2 learners.

opposing opinions with reasons and rebuttals. Ferretti et al.'s (2000) rubric was adapted by many following researchers, such as Nussbaum and Kardash (2005), Nussbaum and Schraw (2007), Chase (2011), and Ong (2013), for the evaluation of persuasive passages.

From the literature review, it was clear that although the researchers created or adjusted the previously used ones, the holistic rubrics' basic components stayed quite similar. Two important things considered were the presence of a position/opinion and supporting data. The other important thing was the strength/soundness of data which supported the positions, etc. Generally, they also expected the writers to consider the opposite ideas (counter positions) together with their rebuttals which can weaken them. Thus, good writers were to stand for a clear position/claim/standpoint, include supporting data, consider alternative views/counterarguments, and try to employ rebuttals. For example, a very convincing passage in Ong's (2013) adjusted rubric includes the following criteria "The essay is overall very convincing. The development of ideas is very good. A stand is provided, with at least three main reasons to support the stand. The reasons are very well-supported and elaborated by examples, reasons, or illustrations. One or two counterarguments are proposed with refutations." (p. 541). Beyond these, one difference between the rubrics which assess the first and the rubrics which assess the L2 was the requirement for the counter position and rebuttal. For example, Qin and Karabacak (2010) and Qin and Lu (2021) included counter position and rebuttal, but they used the modal "may" and stated that the rebuttal is not necessarily required. To illustrate, their level five criteria stated that "The paper may

present reasonable opposing view(s) and also refute the opposing view(s) appropriately, though they are not required.” (Qin & Karabacak, 2010, p. 455).

In the following part, I will share four questions that I considered as I worked on the rubric. They are about the number of levels, the main focus and division, the essential components, the data's soundness, and the rubric's distribution of scores. To note, I will use the term position rather than the other related concepts such as claim, standpoint, stand, and opinion in order to be consistent. Again, the term counter position will be used to refer to the alternative view, opposite view, or counterclaim.

The first question is how many levels must be employed. Most of the rubrics in the literature range from 5 (e.g., Erduran, Simon, & Osborne, 2004; Majidi et al., 2021) to 7 (e.g., Ferretti et al., 2000) levels. And, believing that keeping more levels can increase the precision of the evaluation and the increase in the scores can be more ordinal, I will create a rubric with seven levels.

The second question is what the main focus of the rubric must be. It may be important to remind that in the main writing task of this study, the participants were asked to think about the importance of the needs of clubs and decide how much money to allocate for each of them. The aim would be to convince people/readers by supporting their ideas with reasons and examples. Thus, “whether a passage is overall convincing or not” must be at the core of this holistic rubric.

The third question is what the core components of the rubric must be. As I mentioned above, many persuasive writing researchers gave importance to the existence of a clear position, supporting reasons for that position, the elaboration of the supports (examples and explanations), the recognition of counter positions, and an attempt to

weaken them (e.g., Midgette et al., 2008). In line with this observation, the things to consider can be the existence of a position for each club, counter position(s), rebuttals, and the data which support/weaken/refute these positions. Because the task of this study asked the participants to do so, a well-written passage must reveal the writer's position for all three clubs. In the literature, it is accepted that an explicit statement of viewpoints may increase the level of persuasiveness effect of the texts (O'Keefe, 1997). Thus, the statement of a clear position is important. In our study's context, some participants stated their positions for all three clubs in a single sentence, while many others distributed them across the passage. To illustrate the latter, a participant first mentioned that he would give 5000 liras to the outdoor club. Then, in another part, he stated that he would give 25000 liras to the theater club, and then in another part, 50000 liras to the music club. These kinds of distributions are acceptable. Secondly, the counter position and rebuttal may also be important indicators of a well-written passage. They can signal whether a writer overcomes his/her my-side bias and contributes to the level of persuasiveness. Thus, although they are not necessarily required to get a high score, their existence may contribute to the evaluator's decision. Lastly, it is certain that these positions, counter positions, and rebuttals need the data (reasons, examples, etc.) which can support the writer's position or counter position or can weaken/refute the counter position.

The fourth question is about the soundness of the supporting data. According to some researchers (e.g., Stapleton & Wu, 2015; Zhang, 2018), considering not only the pure existence but the quality of the supporting data is important. It is certain that this is a holistic rubric, and the data will not be assessed individually. However, it is important

to be aware of soundness, for example, whether the data is acceptable, whether it is logically and pragmatically valid, etc. (Eemeren, Henkemaans, & Grootendorst, 2002) to evaluate the development of ideas. In line with this, the rubric will emphasize different levels of overall soundness of the supporting data by inspiring, for example, from Stapleton and Wu's (2015) non-acceptable versus weakly acceptable versus acceptable categorization. Also, if a position is supported, and the support is elaborated well, it must not be weakened by irrelevant information or inconsistencies in the data (Ferretti et al., 2000).

Based on these considerations, a primary trait-scoring guide created by Ferretti et al. (2000) and used/adapted by many following researchers in both L2 (e.g., Chase, 2011; Ferretti et al, 2009) and L2 (Ong, 2013; Qin & Karabacak, 2010) are adjusted to evaluate overall persuasiveness of the passages. Apart from Ferretti et al.'s (2000) original rubric, its adjustment by other researchers, particularly by Ong (2013) and Qin and Karabacak (2010), is considered.

When it comes to the validation of the rubric, the following steps were adapted for finalizing it. After completing the literature review and creating an outline based on my notes, I randomly chose seven passages and, together with one of my colleagues³⁹, analyzed the content/criteria which can contribute to the persuasiveness of the passage. Then, I created a list (position, counter position, etc.) and randomly chose 32 more passages to analyze their content further. Because EFL learners wrote the passages, it was impossible to clearly label the statements in many cases. However, this investigation helped me note down the most important criteria and to revise the holistic

³⁹ Special thanks to Mehtap Güven Çoban for her valuable negotiation.

rubric in several steps. Following this, I used the revised rubric to evaluate the randomly chosen 32 passages, finalized the rubric, and asked a testing member⁴⁰ to check the wording, etc. Then, I evaluated all the passages and revised the rubric when needed. Finally, around ten days later, I evaluated the passages once again and checked the consistency of my scoring. After establishing the validity and reliability of the rubric, a second grader⁴¹ and I, as the researcher, evaluated all the passages using the finalized holistic rubric by following the steps below.

The evaluation of the passages included the following seven steps.

In the first step, before starting to evaluate, as Jacobs et al. (1981) suggested, the first and the second evaluator talked about the aim of the study and the intended use of the results.

The second step was familiarizing with the task of the study, the background of the rubric, the rubric itself, the definition of core components (see above), and this guideline. After reading them carefully, the second rater came together with the researcher to clarify if there were any ambiguities.

At the third step, 16 papers (~10 %) from seemingly different levels were chosen by the researcher who had already evaluated them. Then, they were evaluated by the second rater, as well. It was emphasized that each paper must be evaluated on its own merit considering the criteria in the rubric and following the steps in this guideline. It is important to note that, as Ferretti et al. (2000) stated, the main concern of the rubric was the level of persuasiveness of the passages (their potential to influence the audience's thoughts and behaviors, etc.). Thus, the raters were directed to judge the

⁴⁰ Special thanks to Laura Susan Domrose for her valuable feedback.

⁴¹ Special thanks to Alena Dhahi for her valuable support as a second evaluator.

overall persuasiveness of the passage first. At this initial evaluation, the raters could ask whether the text has the potential to change the readers' mind about the allocation of money. Then, the texts were to be read for a second time to give a more specific score between 0-7. In this second reading, the grade of the passage could be finalized particularly by considering whether the positions about the allocation of money were stated, whether acceptable enough supporting data was provided to persuade the reader, whether supporting data was provided and elaborated via examples, etc., and whether opposing positions were addressed. It may be important to be aware that a highly evaluated text must provide positions for all three clubs and support them with data. Further, the support must be explained and elaborated with examples, comparisons, appeals to authority, etc. Importantly, although it is not necessarily required, it may also consider alternative explanations and provide data that can weaken or refute them. Doing this may increase the score of the passage. Lastly, a good passage mustn't be weakened by inconsistencies in claims or reasons and by irrelevant information.

At the fourth step scores of the independently evaluated papers were compared, and the discrepancies (4 out of 16 passages) were resolved.

Then, at the fifth step, the rest of the papers were evaluated independently. To neutralize the tendency of evaluators to mark lower close to the end of the evaluation period, the passages were divided into two and randomized differently for both evaluators (Jacobs et al., 1981). The general consistency was checked after the second evaluator completed the evaluation of the first half of the passages. This ensured that both evaluators followed the same/similar standards (Jacobs et al., 1981).

The aim of the sixth step was to self-check for consistency. After completing the first cycle, the evaluators ordered the papers from the lowest to highest scores and checked their consistency in the evaluation (scores) by considering the rubric again.

Lastly, at the seventh step, the consistency across the evaluators was checked. The scores from both raters were compared, the discrepancy cases (if two scores are more than one point apart) were marked, and interrater reliability was calculated. I mainly considered Cronbach's alpha to calculate inter-rater reliability. Other than that, as Ferretti and Lewis (2019) did, I also considered interrater agreement within 1 point difference (agreement = agreements/agreements + disagreements). The Cronbach alpha was quite high ($\alpha = .83$), and 85.1% (21 out of 141) of papers were within 1 point difference limit. The scores which had more than one point difference were finalized by negotiating, and then the average of the scores was taken.

5.5.2 Control measurements

5.5.2.1 Coding self-reports in the retrospective questionnaire

As mentioned in the materials section above, the qualitative part of the retrospective questionnaire mainly consisted of task-related background (previous writing experience, etc.) (Appendix J) and writing reflection questions (whether the prompt was understandable, etc.) (Appendix K). To be able to systematically compare the experience of the participants, each answer was categorized and then coded by two (sometimes by three) trained coders⁴². The procedure was briefly shared in the following part (please see Appendix J and Appendix K for the finalized code guides).

⁴² Special thanks to Dr. Gülsen Yılmaz, Onur Keleş, and Oylum Savlak for their valuable support as second and third coders.

I essentially adopted an inductive coding approach for creating a codebook for the retrospective questionnaire data. That is, there was no a priori coding list or manual, but the codes emerged progressively as I investigated the answers given by the participants (Miles, Huberman, & Saldana, 2014). After assigning descriptive codes that summarized the answers' core content (Miles et al., 2014), I created a manual, revised it circularly (coding, going back to the manual, recoding, etc.), and coded the data. Then, I asked one independent rater to code the same data by using the manual and give feedback about it. In the first step, the researcher shared the details related to the writing tasks, the questions asked, the answers obtained from the participants, and the codes and their definitions. After the second coder familiarized himself with the material, a meeting was organized, all the procedure was reviewed, and any ambiguities were clarified. Following that, around 10% of the data was coded independently, and the consistency was checked. In this and the following steps, the manual was revised progressively when needed, and the data was independently recoded by using the new version of the manual. Following that, I checked interrater reliability to see if a reasonable level of reliability was reached.

With respect to interrater reliability, because different approaches were reported in the literature, one of the issues was to decide how to assess it for this mostly categorical data. The literature review revealed that the assessment can be done by simply checking percent agreement among the coders or by calculating Cohen's Kappa (κ), which takes the chance factor into account (random guesses made by a coder when he/she was not sure) and which can range between -1 and +1. Its values can be

interpreted as the following: 0-.20 (no agreement, 0-4% reliability), .21-.39 (minimal agreement, 4-15% reliability), .40-.59 (weak agreement, 15-35% reliability), .60-.79 (moderate agreement, 35-63% reliability), .80-.90 (strong agreement, 64-81% reliability), above .90 (almost perfect agreement, 82-100% reliability) (McHugh, 2012).

When compared with percent agreement, κ inter-rater reliability analysis is suggested to be more robust. However, as McHugh (2012) reported, both have strengths and limitations. Percent agreement⁴³ is easy to interpret and calculate but is not robust for the chance factor. Still, it can be relied on, particularly in a context where researchers are less likely to make guesses during coding. κ considers the chance factor, but as some researchers (e.g., Syed & Nelson, 2015) reported, it is a conservative index of reliability: Its formula ($\kappa = (\text{observed} - \text{expected chance}) / (1 - \text{expected chance})$) is not only corrected for a chance, but completely removed it from the index. And because of the properties of its formula, the high κ values cannot be achieved in the context of some studies, and low κ does not always mean low interrater reliability (e.g., Bajpai, Bajpai, & Chaturvedi, 2015). Moreover, its values are not easily interpreted, and some of its assumptions were not supported well. Still, it can be useful, particularly if most of the observations fall into a single category and inflate the percent agreement scores (Stemler, 2004). Considering these, McHugh (2012) recommends calculating both percent agreement and κ scores to assess interrater reliability. This can be more informative because it is evident that a $\kappa = .64$ with 95% agreement level should not be interpreted in the same way as a $\kappa = .64$ with 95% agreement (Syed & Nelson, 2015).

⁴³ It is also more flexible in that it can be used in different types (nominal, ordinal, and continuous) of data (Syed & Nelson, 2015).

When it comes to the benchmark of the accepted reliability, although there are different reports, κ values between .61-.80 are categorized as *substantial* and κ values between .81 and 1 are accepted as *almost perfect* (Landis & Koch, 1977) but generally $\kappa = .70$ and above is considered as an appropriate level of agreement between the raters (Syed & Nelson, 2015) and many researchers reported κ values around .70 (e.g., Hallgren, 2012; Hu & Cao, 2015) as a valid reliability score. Regarding percentage agreement, although some researchers accept a lower percentage depending on the coding complexity (e.g., Belotto, 2018; Campbell, Quincy, Osserman, & Pedersen, 2013), %80 percent agreement is considered a benchmark by many researchers. In this study, I primarily considered κ values and expected them to be higher than .7 but also reported the percentage values.

In accordance with the suggestions in the literature (e.g., Hruschka et al., 2004; Krippendorff, 2004), there was an iterative process of coding, followed by a reliability assessment, and by the modification of the codebook, such as giving more details or reduction/combination of the codes and recoding. The whole data was coded individually by both coders in each circle, and the circles continued until the κ interrater reliability value reached at least .70. The following table, which includes the translated questions, summarizes Cohen's κ and percentage agreement values reached before discussing the discrepancy and gives a reference to the related guideline (Table 8 and Table 9).

Table 8. Retrospective Writer Background Questions and the Interrater Reliability Values

	Writer Background Questions	Cohen's κ	Agreement %	Coding Guideline Details
Q1	<i>Did you take English writing courses before coming to the university, including primary and high school? If you did, can you give details about the course content, materials, assignments, feedback, etc.?</i>	0.91	96.40 %	Appendix J Table J1
Q2	<i>Can you please evaluate your English writing performance in general?</i>	0.72 (0.74)	80 % (84.3 %)	Appendix J Table J2 Table J3
Q3	<i>Did you take Turkish writing courses before coming to the university, including primary and high school? If you did, can you give details about the course content, materials, assignments, feedback, etc.?</i>	0.73	85.70 %	Appendix J Table J4
Q4	<i>Can you please evaluate your Turkish writing performance in general?</i>	0.80 (0.76)	89.1 % (87.8 %)	Appendix J Table J5 Table J6
Q5	<i>Do you usually think in Turkish or in English while writing? Can you give details?</i>	0.72	87.10 %	Appendix J Table J7
Q6	<i>What kind of writing activities do you do in or out of class? Can you give details?</i>	0.80	85 %	Appendix J Table J8
Q7	<i>Do you have a special interest in music, theater, or outdoor sports? Can you give details?</i>	0.84	85.70 %	Appendix J Table J9
Q8	<i>Are you a member of music, theater, or outdoor sports clubs? If you are not, do you plan to be? Can you give details?)</i>	0.77 (0.81)	81.4 % (94.3 %)	Appendix J Table J10

Note: If the coding included a second level/step (a subcategory such as the source of writing difficulty for the second question, see Appendix J), its value is given in parentheses. All P s > .001.

Table 9. Retrospective Writing Reflection Questions and the Interrater Reliability Values

	Writing Reflection Questions	Cohen's κ	Agreement %	Coding Guideline Details
Q1	<i>Did you consider the audience (rector and club members) perspective while writing? If you did, did it influence your writing? How? Can you give details?</i>	0.71 (0.72)	78.8 % (79.6 %)	Appendix K Table K1 Table K2
Q2	<i>Was the writing question easy to understand? Can you give details?</i>	0.90	97.10 %	Appendix K Table K3
Q3	<i>Did you experience difficulty while writing the text? Can you give details?</i>	0.86 (0.80)	90.7 % (83 %)	Appendix K Table K4 Table K5
Q4	<i>Did you have a chance to review the text after writing it? Can you give details?</i>	0.80	91.40 %	Appendix K Table K6
Q5	<i>How do you think this text you wrote will be evaluated? Can you give details?</i>	0.76	83.60 %	Appendix K Table K7

Note: If the coding included a second level/step (a subcategory such as the influence of considering the audience on writing for the first question, see Appendix K), its value is given in parentheses. All P s > .001.

After reaching the plausible intercoder ($\kappa = .70$) reliability level, the coders discussed the discrepancies (Campbell et al., 2013) to reach a 100 % agreement. If they still cannot, a third coder was included to help to recode and reach a consensus about the controversial items.

5.5.2.2 Coding strategy use in EF tasks

After collecting the EF data, an online survey was prepared, and the participants were asked to report if they used any strategies which were not part of the task instructions but could influence the obtained results. Similar to the retrospective questionnaire, I

adopted an inductive approach for coding the reported EF strategy use. The codes emerged progressively as I investigated the answers given by the participants (Miles et al., 2014). Additionally, I checked the literature to see if my initial codes were plausible. The comparisons revealed that my observations were in accordance, for example, with Dunlosky and Kane (2007) and Davelaar, Hannon, and Richards (2005), who reported strategies such as developing mental images and rehearsing.

The answers were coded by the researcher and one independent rater⁴⁴. In the first step, the researcher shared the details related to the EF tasks, the questions asked, and the answers obtained from the participants, as well as codes and their definitions. After the second coder familiarized himself with this material, a meeting was organized, all the procedure was reviewed, and any ambiguities were clarified together. Following that, around 10% of the data was coded independently, and the consistency was checked. Then, the whole data was coded again independently, the final codes were compared, and interrater reliability (Cohen's κ) was evaluated. If the interrater reliability was not around .70 or higher, the coding guide was revised, and the coding was done once again. After reaching a plausible consensus, the researcher and second coder came together, and the discrepancies were resolved by discussion. See Appendix L for more details about the codes and the following Table 10 for the interrater reliability values.

⁴⁴ I thank Onur Keleş for his valuable support in coding.

Table 10. Questions for the Strategy Use in EF Tasks and the Interrater Reliability Values

	Strategy Use Questions	Cohen's κ	Agreement %	Coding Guideline Details
Q1	<i>Did you use any strategies as you completed the attention task during which you chose the direction of the arrows (left or right)? If you used, can you give details?</i>	0.71	92.6 %	Appendix L Table L1
Q2	<i>Did you use any strategies as you completed the attention task while trying to catch the O or Q letters by moving your eyes to the left or right when an asterisk appeared on the screen? If you used, can you give details?</i>	0.85	91.9 %	Appendix L Table L2
Q3	<i>Did you use any strategies as you completed the memory task during which you tried to evaluate whether the presented figures were symmetrical while simultaneously trying to keep the location of the squares in your mind? If you used, can you give details?</i>	0.75	90.4 %	Appendix L Table L3
Q4	<i>Did you use any strategies as you completed the memory task while trying to solve mathematical problems while simultaneously trying to keep the letters in your mind? If you used, can you give details?</i>	0.81	87.4 %	Appendix L Table L4

All $P_s > .001$.

5.5.3 Manipulation check measurements

In this study, I collected four different types of data which helped me understand whether the experimental task successfully oriented the writer to the audience before and while writing. Two of them based on self-reports collected via the retrospective questionnaire. The first one was the five-point scale which asked the participants

whether they tried to understand the audience while reading the vignettes during the main writing task and the second one was a retrospective question which asked whether they considered the audience while writing. The details about these materials were given above. Beyond these, I considered brainstorming and PT activity notes taken during task processing (before starting to write) and also coded the main writing passages to check the manipulation from the researcher's perspective. Further details about these latter two manipulation check measurements will be given below.

5.5.4 Coding PT and brainstorming activity notes for manipulation check

As mentioned before, PT is defined as “the active cognitive process of imagining the world from another’s vantage point or imagining oneself in another’s shoes to understand their visual viewpoint, thoughts, motivations, intentions, and/or emotions” (Ku et al., 2015, p. 94-95). Following this definition and considering the items in Davis (1980) scale, I created a guideline and coded the notes the participants taken during the 3 minutes PT or brain storming activities (after reading the vignettes but before writing the main essays). This assessment added another dimension for understanding whether the participants in the experimental group were engaged with the audience's perspective more when they were asked to reflect on the vignettes by taking the audience's perspective (see Appendix M for the guideline details).

5.5.5 Coding Interactional Markers in the Main Passages for Manipulation Check

One of the most important checks within this research's context may be related to whether the experimental audience-oriented task made the participants maintain the

audience in their minds while writing their main passages more. Among many others, one indicator of a writer's attempt to consider the audience while writing may be the employment of engagement markers (EM), which "represent a reflective awareness of self, text, and audience" (Hyland, 2010, p.134) and signal reader awareness in a written passage (Hyland, 2005).

These markers (e.g., inclusive we, you, the verb consider, etc.) which aim to build a direct relationship with the readers by referring (Hyland, 2005) and including them as the participants of the text (Hyland, 2010) can be the explicit indicators of the reader-writer interaction (Hyland & Jiang, 2016; Zou & Hyland, 2020). That is, as Fang and Zhuang (2022) suggested "The use of EM can explicitly reveal how the author cares for and interacts with the audience" (p. 87) and can be used to explore the attempt of a writer to interact with the audience in an explicit/concrete way.

On the opposite side, self-mention (SM) (I, me, my, exclusive we, etc.) markers which aim to give an explicit reference to the writer (Hyland, 2005), can be the most potent self/author representation (Hyland, 2001; Ivanic, 1998), and because they emphasize the presence of the author (Hyland, 2010) sometimes their absence, rather than presence, can be considered as an indication of audience awareness. In other words, these self-appealing markers can be associated with the relative absence of an effort to engage the audience (Midgette et al., 2008) and can be used to check state PT manipulation (Galinsky et al., 2005).

Based on these, in this study's context, I considered these markers as one of the instruments for the manipulation check (maintaining audience while writing) and coded

the frequency of their employment in the main writing passage. See Appendix N for the guideline details.

5.6 Ethical concerns

5.6.1 The protection of the participant identity

In order to keep the data confidential and avoid identity and data match, I randomly gave each participant a three-digit number at the beginning of the experiment (after collecting consent forms).

The codes were not matched with the identity (name) of the participants except in three cases. First, they were used to differentiate the groups from each other (e.g., experimental versus control). Second, they were used to contact some of the participants to invite them for post-experimental data collection (e.g., EF). Third, they were used to inform the participants about the study's results. Apart from these, all the experimental material (background questionnaire, diagnostic writing task, main writing tasks, retrospective questionnaire, EF tasks, and PT scale) was processed by using randomly assigned codes and these codes were not matched with the names of the participants. All the analyses were conducted as a group, and if a specific case (e.g., a quotation) needed to be reported, only the participant's code was used.

5.6.2 The protection of the participant data

The printed materials of the study were protected in a locked file cabinet at the researcher's office during the data collection. After that, all the materials were scanned, the printer versions were destroyed by putting them in a cross-cut shredder, and the

scanned versions were protected in a password-protected file in the researcher's computer and hard disk.

The scanned essays (both diagnostic and experimental main writing) were transcribed, and all the data (including background, retrospective questionnaire, AC, WM, and trait-PT) were entered into Excel files. The Excel file containing participant names and codes was kept in a separate folder. Only this file included the participants' names and codes; as with all the experimental materials, it is also protected in the researcher's external hard disc. The Excel and SPSS files, which included experimental data for the analysis and other experimental materials, are saved in a separate folder.

The project conductor also had full access to all these files, and upon her request, the researcher could share the files with a passcode. The third team members (e.g., the second evaluators) had access to some data (e.g., the essays) of the participants they need to use but not the background information or any other data.

CHAPTER 6

RESULTS

This study adopted a cognitive approach, considered EF capability and trait PT tendency of the participants, and investigated the efficiency of directing the attention of novice upper-intermediate level Turkish EFL writers to their audiences via an audience-oriented experimental task. The participants were assigned either to an audience-oriented experimental or non-audience-oriented task control condition. After processing the task components, they were asked to write a convincing passage. Besides the experimental manipulation, EF capability and trait PT tendency functioned as independent variables. The dependent variables were the overall writing quality and the level of persuasiveness of the passages.

In addition to these, some background data (e.g., the tendency to think in Turkish while writing in English, having taken English writing courses, etc.) were collected to control the similarity of the groups, some reflection data (e.g., the perceived difficulty of the prompt) were collected to check whether, for example, the task was understood, and some manipulation check data (e.g., self-reports about taking the audience perspective) were collected to see if the manipulation worked and the experimental group were engaged with the audience while writing more. Hence, although I may use them to support the main experimental data and make conclusions, their primary aim is to control whether there is an eye-catching group difference which needs to be dealt with and to check whether the manipulation worked. Other than that, they are not part of the main questions or analyses.

This chapter will comprise preliminary, main, and additional analyses sections. The preliminary section will start with data preparation, the participant background comparison, and the investigation of participant reflections on the experimental task. The section will be complete with the manipulation check. Then, the main analyses section will focus on and try to answer the main questions of the study. Finally, the additional analyses section will consider two analyses that can help get further insight into the data. If not otherwise stated, the inferential analyses will be conducted as two-tailed, and the missing data will be dealt listwise.

6.1 Data preparation and preliminary analyses

In this section, I will first organize, check, and compare the background data of the experimental and control groups. I expect both groups to be similar with respect to their EF capabilities, trait PT, initial proficiency, present vocabulary, present grammar, present reading comprehension, and general writing performance as well as their interest and involvement in the student clubs and related activities. Following that, the groups' retrospective reflections will be compared to get insight into the patterns revealed by the groups' self-reports about the task. Finally, before starting the full-scale main analyses, four types of data will be investigated to check if the main experimental manipulation was successful. I expect the experimental and control tasks to influence the participants in different ways and to orient the experimental group to their readers more.

6.1.1 Data preparation and background comparison

After collecting the paper-based data in the classroom environment, I transcribed all the written essays into a Word document without making any corrections and checked the similarity rates in Turnitin to see if the participants directly copied anything from the vignettes. The similarity rate was 0 for all the participants⁴⁵. The rest of the data were entered into an Excel processor and, if needed, coded and re-coded for the analyses. To note, only the data of the subjects who completed the main writing, EF tasks, and trait PT scale ($n = 133$) were included.

I checked the whole data to see if there were any data entry problems or some unnatural cases, such as repeatedly pressing the same (left or right) button or responding too quickly (faster than 200 ms) or too slowly (slower than 1000 ms) in the flanker task (see below). Again, I checked for any missing data and took the necessary measures. For example, I excluded the initial proficiency scores of the participants if they did not complete all sections of the exam. Beyond this, to check the background of the participants across the whole group and as a part of data cleaning, I also calculated z-scores for continuous data (Tabachnick & Fidell, 2013) and visually inspected the distributions (e.g., histograms with a normal curve) for detecting any extreme outliers. The aim was to see whether there were any background data (language-related variables, EF tasks, and trait PT scale) which were extremely different from the rest of the participants. I would treat/exclude that data as an outlier if its distribution seemed

⁴⁵ This is not something surprising within the context of our school where the students use some source texts for writing, and they were punished if the similarity exceeds some limits.

unnatural and the data's *z-value* was lower than -3.29 or higher than 3.29 (Tabachnick & Fidell, 2013).

My investigation revealed an extreme *z-score* (participant 303, *z-score*: 5.50) in the flanker task. And because the following visual investigation confirmed that this could be unnatural, I removed it from the data. Other than this, although there were a few cases that exceeded the *z-score* of 3.29, the total number of them was less than 1%, the shapes of the data were close to a normal distribution, and the data seemed natural. Thus, they were not excluded (Gadermann, Schonert-Reichl, & Zumbo, 2010).

The experimental data, including overall writing quality and the level of persuasiveness of the written passages were also checked as part of data cleaning and as part of inferential tests' general assumptions (e.g., normality and having extreme outliers). Their distributions were quite natural.

Now, I will focus on EF capability, trait PT ability, and language and topic related (club membership, etc.) background of the participants.

6.1.1.1 Executive function tasks

As reported in the methodology section, there were four tasks that aimed to assess the EF performance of the participants. Two of them, flanker and antisaccade, were to assess the AC capacity, and the other two, foster operational span and symmetrical span, were to assess WM capacity.

The reason for employing four tasks was to collect the data with sufficient variability. This would be in accordance with the writing models, which include both AC and WM as the limited cognitive processes which create a foundation for writing

processes (e.g., Hayer & Berninger, 2014; Kim & Park, 2019) and with the models which approach EF as a unitary concept (Engle, 2002). Otherwise, if only one task were applied, it would possibly be more prone to noise, the study could miss some informative data⁴⁶ and have difficulty observing the real EF performance of the participants.

As applied/recommended by some researchers on the field (e.g., Crane et al., 2008; Gibbons, et al., 2012; Godefroy et al., 2014; Godinho, Maruta, Borbinha, & Pavao Martins, 2022; Roberts et al., 2022), one of the ways to approach this kind of a data is to create composite scores (Kolovelonis, Pesce, & Goudas, 2022; Moreau & Wiebels, 2021; Rızaoğlu & Gürel, 2022). Following this, I created three composite scores by calculating z-scores for each EF task and merging them by taking their averages. AC composite scores consisted of the flanker and antisaccade data, WM composite scores consisted of symmetry span and foster operational span data, and overall EF composite scores consisted of the average z-scores of all four tasks' data. In the following parts, I will describe these composite variables after cleaning and preparing the raw AC and WM data.

6.1.1.1.1 Flanker attention control task

The participants' speed in reporting the direction of the congruently or incongruently presented arrows was the main data of the flanker task. To clean it, first, the responses

⁴⁶ For example, if Baddeley's (2000) operationalization is more explanatory, collecting only spatial location-based WM data would refer to visuospatial sketchpad and collecting only phonology based data would refer to phonological loop components. This could be the case, because for instance, it was revealed that a reading span WM task may be associated with verbal rather than visuospatial abilities. In contrast, a spatial span WM task may be associated with visuospatial rather than verbal abilities (e.g., Shah & Miyake, 1996).

faster than 200 msec and slower than 1000 msec, and then all inaccurate responses were removed. After that, the average reaction time of the participants to congruent and incongruent stimuli was calculated separately, and the reaction time to congruent stimuli was subtracted from the reaction time to incongruent stimuli. Finally, I conducted a between-subject t-test (strategy use versus no strategy use) and univariate ANOVA (for the subcategories of strategy use). I saw that the reported strategy uses (see Appendix L) did not differentially influence the obtained results ($p > .1$).

The overall mean of the response time differences between incongruent and congruent trials was 37.77 msec ($SD = 18.62$), with a range from .139 to 121.35 msec. And the between-subject t-test comparison revealed no difference between the experimental ($M = 36.68$, $SD = 19.78$) and control ($M = 38.83$, $SD = 17.51$) groups, $t(130) = .661$, $p = .510$.

6.1.1.1.2 Antisaccade attention control task

The accuracy of detecting the letters (Q or O), which were shortly presented in the opposite direction of the initially presented cue (an asterisk), was the main data of the antisaccade task. As part of data cleaning, the unnatural responses faster than 100 msec (79 trials, less than 1%, in total) were removed. The final mean reaction time range was between 330 msec and 1207 msec. Because the main concern of this task was the accuracy rather than the speed, and because there were no extremely slow cases, I did not remove three relatively slower (> 1000 msec) responses. Finally, a between-subject t-test analysis and univariate ANOVA revealed no effect of the strategy use (see Appendix L) ($p > .1$).

The mean accuracy of the participants was 51.72 ($SD = 8.999$), with a range from 33 to 69, and a between-subject t-test result revealed no difference between the experimental ($M = 51.64$, $SD = 9.313$) and control ($M = 51.81$, $SD = 8.749$) groups, $t(131) = .108$, $p = .914$.

6.1.1.1.3 Symmetry span working memory task

As described in the materials section, this task included a symmetry processing control within itself. The participants were asked to reach at least 85% accuracy in their symmetry evaluation. The aim was to keep them engaged and try to prevent them from using some strategies, such as rehearsing the target location of the boxes. Thus, I checked the data to clean those who got low scores from the symmetry evaluation. However, because I followed the participants carefully to see if they focused on evaluating the symmetry of the figures, I determined 80%, rather than 85%, as a criterion point. This resulted in excluding 5 participants who got lower than 80%.

Finally, I conducted a between-subject t-test to see if the reported strategy use (see Appendix L) influenced the obtained results. The analysis revealed an effect of the strategy use, $t(118) = 1.991$, $p = .049$. To understand the source of this effect, I further analyzed the data by conducting a one-way ANOVA with post hoc Bonferroni corrected tests. The main comparison was between no strategy use and the strategy use categories (coding/grouping, rehearsal, visualization, and other strategy use). The results revealed that the source of the effect, $F(5, 122) = 3.001$, $p = .014$, was the difference between no strategy use ($M = 27.82$, $SD = 7.359$) and other strategy use ($M = 15.25$, $SD = 6.966$), $p = .001$. However, before making a final decision, I also checked the participant reports

to make a final decision because there were only four participants in the other strategy use group (still, homogeneity of variances assumption was confirmed). Seeing that their strategies were illegal and could potentially interrupt the appropriate data collection (e.g., one of them said that “he marked the locations by using his hands ...”), I decided to remove those four participants from the symmetry WM span data.

After cleaning, the mean partial span score of the participants was 27.42 ($SD = 7.19$), ranging from 33 to 69. And, between subject t-test analysis revealed no difference between the experimental ($M = 28.03$, $SD = 7.094$) and control ($M = 26.81$, $SD = 7.290$) groups, $t(122) = .949$, $p = .345$.

6.1.1.1.4 Foster operational span working memory task

Similar to symmetry span, the foster operational span task included a processing control, and basic math calculation, within itself. The participants were asked to reach at least 85% accuracy in their basic match calculation to keep them engaged and prevent them from using some strategies, such as rehearsing the target letters as much as possible. Again, I observed all the participants and checked if they did their best to complete the task accordingly. Thus, I determined 80%, rather than 85%, as a criterion point and cleaned those who got lower than 80% in the math calculation part. This resulted in the exclusion of 3 participants. Finally, I conducted a between-subject t-test and a univariate ANOVA to see if the reported strategy use (see Appendix L) influenced the obtained results. The analysis revealed no differential effects of the strategy use ($p > .1$).

After cleaning, the mean partial span score was 37.42 ($SD = 7.518$), ranging from 14 to 50. And a between-subject t-test analysis revealed no difference between the

experimental ($M = 37.53$, $SD = 7.368$) and control ($M = 37.32$, $SD = 7.716$) groups, $t(128) = .161$, $p = .872$.

6.1.1.1.5 Composite EF performance scores

After cleaning and preparing the EF data for each task, I created three composite scores of the participants by calculating z-scores and then merging them by taking averages. To note, because minus z-scores of flanker data meant better performance (faster reaction to incongruent trials), I multiplied them by -1 before merging so that it would be in line with the others. The finalized composite scores were AC composite (flanker and antisaccade data), WM span composite (symmetry span and foster operational span), and Overall EF composite (the average of four tasks' z-scores). I will mainly focus on EF composite scores during the analysis and consider AC and WM composite scores (or other subcomponents) only if necessary.

Following data preparation, I compared the composite scores across the experimental and control groups. Between-subject t-test results revealed no significant performance difference in AC composite, $t(130) = .431$, $p = .667$, WM composite, $t(120) = .716$, $p = .475$, and overall EF composite scores, $t(118) = .585$, $p = .56$.

6.1.1.1.6 The split of low versus high executive function capability

To determine the groups with relatively lower and higher EF capacity and to compare their writing performance across the experimental conditions, I split⁴⁷ the AC

⁴⁷ Splitting data was done in different ways depending on the purpose and the context of the study and used by many researchers in the area (Kane, Bleckley, Conway, & Engle, 2001). Still, I also conducted

composite, WM composite, and EF composite scores from the median. Depending on the number of participants, the median was added either to the low or high-capacity groups (Table 11). To repeat once again, the overall EF composite score will be the primary consideration in the analysis.

Table 11. The Median Z-Values of AC, WM, and EF

EF Tasks	Median
Attention Composite	0.02
WM Span Composite	0.07
Overall EF Composite	0.08

6.1.1.2 Trait perspective taking scale

As reported in the methodology section, there were seven items in the trait PT Scale, and two of those items were reversed. The data of those two reversed items (the first and the fourth items) were re-calculated, and an interitem reliability analysis was conducted. The obtained Cronbach Alpha for the scale was $\alpha = .70$. Seeing that the reliability was within an acceptable range, the average of the items for each participant was taken. The mean was 3.656 ($SD = .665$), ranging from 1.43 to 5.00.

Following this, a between-subject t-test analysis was conducted to compare the groups and to be sure that the experimental and control groups were similar. The results revealed that the experimental ($M = 3.59$, $SD = .704$) and control ($M = 3.72$, $SD = .625$)

additional analyses such as regression and correlations with continuous data which enriched the analyses by adding further dimensions.

groups were not significantly different in their overall trait PT tendency, $t(131) = 1.088, p = .279$.

Then, similar to EF tasks, I split trait PT data from the median (3.71). The data above the median were categorized as relatively high, while the rest were categorized as relatively low perspective takers. After splitting, the mean value of the relatively higher group was 4.20 ($SD = .31$), while the mean value of the relatively lower group was 3.17 ($SD = .51$).

6.1.1.3 Language background

6.1.1.3.1 Self-reported writing background

After writing their main passages, some writing-related background responses were collected from the participants and coded for objective comparisons. The aim was to see whether the experimental and control groups were similar to each other. Please see the Retrospective Questionnaire Coding section above and Appendix J below for this and the following questions' coding procedure.

Most participants reported that they did not take English writing lessons before the preparatory school, and the distribution of those who took English writing and those who did not take English writing lessons were similar across the groups. This was also the case for Turkish writing lessons (Table 12).

Table 12. The Distribution of those who Took and did not Take English and Turkish Writing Lessons Before the Preparatory School

Previous Writing Lessons		Experimental	Control
In English	Yes	16	17
	No	47	48
	Other	3	2
In Turkish	Yes	36	40
	No	26	21
	Other	4	6

The participants were also asked to evaluate their English and Turkish writing performance. The answers were mainly coded at three (almost no difficulty, somewhat difficulty, and (significant) difficulty) levels (Table 13). The frequency values indicated that the groups were similar to each other in this respect, as well.

Table 13. The Distribution of the Reported Difficulty in General English and Turkish Writing

Difficulty in General Writing		Experimental	Control
In English	(Almost) No Difficulty (1)	9	7
	Somewhat Difficult (2)	28	30
	(Significant) Difficulty (3)	16	17
	Other (4)	3	2
	Irrelevant Answer (5)	9	11
In Turkish	(Almost) No Difficulty (1)	44	37
	Somewhat Difficult (2)	9	8
	(Significant) Difficulty (3)	3	13
	Other (4)	6	6
	Irrelevant Answer (5)	2	2

The responses given by the participants who experienced difficulty were further coded to check the source of the difficulty. Again, the distribution of the participants revealed a similar pattern across the groups (Table 14).

Table 14. The Distribution of the Coded Sources of Reported Difficulties in English and Turkish Writing

The Source	Experimental		Control	
	English	Turkish	English	Turkish
No Difficulty/Irrelevant	19	46	18	40
General Language (including vocabulary and grammar)	5	-	6	-
Grammar	3	-	5	1
Vocabulary	3	-	3	-
General Writing Proficiency	25	9	25	12
Content/Background Gap	1	1	-	-
Content Organization	-	-	4	2
Planning	1	-	-	-
Number of Words	2	1	1	-
Duration	-	1	1	-
Translation from Turkish	4	-	1	-
Mechanics (E.g., Punctuation)	-	-	-	1
Other	2	6	3	10

Another question was whether the participants thought in English or in Turkish while writing in general. Both groups reported that thinking in Turkish was dominant when they wrote something in English (Table 15). This observation eliminated the possibility that one of the groups may asymmetrically depend on translation from the

native language, which can create an extra load and reduce the overall writing quality (Weijen et al., 2009).

Table 15. The Distribution of Reported Thinking Language while Writing in English

Thinking in	Experimental	Control
Turkish	51	48
English	9	9
Both English and Turkish	4	8
Other	1	2

Because there is a possibility that thinking in Turkish may create an extra load (Zimmermann, 2000), I split the experimental and control groups' data and checked whether thinking in Turkish versus thinking in English had an effect on the overall writing quality or the level of persuasiveness. Although the overall writing quality performance of those thinking in English ($n = 9$, *Mean Rank*: 37.61) was higher than those of thinking in Turkish while writing in the experimental group ($n = 51$, *Mean Rank*: 29.25), the *Mann-Whitney U* test results revealed no significant differences between them, $U = 165.5$, $Z = -1.326$, $p = .185$. The pattern was similar in the control group. Again, although the overall writing quality performance of those thinking in English was higher ($n = 9$, *Mean Rank*: 33.28) than those of thinking in Turkish ($n = 48$, *Mean Rank*: 28.20) while writing in the control group, the *Mann-Whitney U* test results revealed no significant differences between them, $U = 177.5$, $Z = -.843$, $p = .399$. The obtained results were similar when I considered the level of persuasiveness, as well (All $P_s > .1$). Hence, the groups were similar to each other.

In addition, although this was not the main aim of this study, to get more insight into the influences of thinking in English within the context of the experimental writing, I invited the participants to respond to an additional online question. The English translation of the question was:

Please go back to the text and time you wrote when you participated in the experiment. When writing the text in English, did you tend to think in Turkish first and then write in English? If so, did doing this make your job easier or harder? How? Can you give details?

The investigation revealed 24 answers (14 of them were from the experimental group), and three of these participants reported that they directly used English while writing. Apart from that, two responses were not clear to make a conclusion, but 19 participants translated their ideas from Turkish into English. Importantly, only five of those 19 participants who tended to translate their ideas from Turkish reported difficulty because of this. For those who shared further details, the main reasons seem to be the difficulty in expressing their translated ideas in English, keeping their originality, and expressing them more simply (Participant 592)⁴⁸, or forgetting their ideas (Participant 509). Those five participants were from the experimental group. Beyond this, as I reported above, most participants (equally distributed across the experimental and control groups) reported that thinking in Turkish first and then translating the ideas into English was helpful for several reasons. The most eye-catching reasons were: it was not easy to retrieve the ideas in English (Participant 513), it helped because the context of the topic was Turkish (the clubs were in Turkey) (Participant 425), helped me think

⁴⁸ Because some participants remember their initial participant numbers, I used their random evaluation codes.

more deeply (Participant 463), helped me think better and analyze the text better (Participant 949), helped me create better sentences in Turkish and translate them into English (Participant 566), and helped me write more fluently (Participant 980). Two of those students found this strategy beneficial but also reported some difficulties resulting from adopting this strategy. One said that although she lost some time and experienced difficulty, it made her feel more comfortable and helped her create more accurate sentences in English (Participant 845), and another student said that although sometimes thinking complex sentences and writing them in English was confusing, she wrote more easily (Participant 468).

Lastly, the participants were also asked about their writing activities during and after the lessons. The distribution (Table 16) indicated that more participants in the experimental group seemed to do activities related to schoolwork; however, univariate ANOVA comparisons revealed that this did not have differential effects on overall writing quality or the level of persuasiveness (All P s > .1).

Table 16. The Distribution of the Writing Activities Done during or after the Lessons

	Experimental	Control
None/Irrelevant	18	26
Personal Writing	17	15
Self (Exam) Study	11	10
Related to School Work	19	12
Other	1	4

6.1.1.3.2 Assessed language background

As mentioned in the methodology section, I collected the diagnostic writing data to familiarize the students with the main writing task and to check whether their general writing performance was similar across the groups. Additionally, with the same purpose in mind, I considered their initial language proficiency and the current grammar, vocabulary, and reading comprehension scores that I collected from the testing unit of the school.

The average grade of the diagnostic writing scores for the whole sample ($n = 133$) was 70.586 ($SD = 8.47$)⁴⁹. And when I compared them, between-subject t-test, results revealed no difference between the experimental ($M = 71.879$, $SD = 8.2$) and control ($M = 69.313$, $SD = 8.6$) groups, $t(131) = 1.760$, $p = .081$.

The vocabulary and grammar exam scores of the students taken from the testing unit of the school did not also reveal any significant differences between the groups. The grammar test average of the participants was 9.86 ($SD = 3.033$), ranging from 1 to 17. And the comparison of the groups revealed that the difference between experimental ($M = 9.88$, $SD = 3.14$) and control ($M = 9.84$, $SD = 2.95$) groups, $t(131) = .81$, $p = .935$, was not significant. When it comes to the vocabulary test average of the participants, it was 12.82 ($SD = 3.155$) with a range from 5 to 20. Again, between subject t-test results revealed no statistically significant difference between experimental ($M = 12.50$, $SD = 3.26$) and control ($M = 13.13$, $SD = 3.04$) groups, $t(131) = 1.161$, $p = .248$.

⁴⁹ Because this mean value is within 65 and 73 range, it may be suggested that the overall writing proficiency of the group was high-intermediate according to the guideline of Jacobs et al. (1981).

I also checked the reading exam scores of students. The average grade of the sample ($n = 132$) was 7.742 ($SD = 1.98$), ranging from 3 to 12. Once again, the comparison across groups revealed no difference between the experimental ($M = 8.03$, $SD = 1.84$) and control ($M = 7.46$, $SD = 2.09$) groups, $t(130) = 1.679$, $p = .095$.

The last instrument I used to check the language background of the participants was the proficiency test the participants took at the beginning of the academic year. The participants' mean score for this test was 46.68 ($SD = 12.71$), with a range from 12 to 91. And again, the comparison of the groups revealed no significant difference between experimental ($M = 45.82$, $SD = 11.698$) and control ($M = 47.51$, $SD = 13.663$) groups, $t(122) = .738$, $p = .462$.

On the basis of the data presented and the analyses conducted above, I can suggest that the groups were similar to each other with respect to their overall language background.

6.1.1.4 Music, theatre and outdoor related activity background

One thing that could affect the experimental manipulation was whether the participants had a special interest in the clubs presented in the tasks. I asked two questions to check this background. One of them was related to the general interest of the participants in the activities mentioned in the vignettes, and the other was about the special interest in the student clubs (becoming a member, etc.). The answers to these questions were coded under nine different categories (Please see the Retrospective Questionnaire Coding section above and Appendix J below for this and the following questions' coding procedure).

As can be seen below (Table 17), particularly when the total number of those interested in one of these clubs, two of these clubs, and three of these clubs are considered, there does not seem to be an important difference between the groups. Moreover, thinking that the topic familiarity (e.g., being active in club-related activities) may influence writing performance, I conducted a series of univariate ANOVA analyses. For example, I regrouped the participants as active in one, two, or three clubs and investigated their interaction with experimental versus control tasks. The dependent variables were the overall writing quality and the level of persuasiveness. The analyses revealed no significant results (All P s > .1).

Table 17. The Distribution of the Reported Interest in Music, Theatre, and/Outdoor Sports Activities and Clubs across the Groups

	General Interest (Finding the Related Activities Interesting)		Club Interest (Active Member/Planning to be Active Member/Actively Did/Doing Something Related)	
	Experimental	Control	Experimental	Control
No Interest/Activity	4	7	20	19
Music	18	11	13	7
Theatre	10	8	9	10
Outdoor	7	15	11	17
Music & Theatre	8	4	5	2
Music & Outdoor	9	9	2	3
Theatre & Outdoor	1	2	1	-
Music & Theatre & Outdoor	7	11	1	4
Not Clear/Other	1	-	3	5

6.1.2 Retrospective reflection on main writing processes

Again, as mentioned above, some retrospective reflection questions were asked after the participants completed the main writings. These questions were also coded to be able to, for example, check the perceived difficulty of the task and also compare the groups more systematically. Please see the Retrospective Questionnaire Coding section above and Appendix K for this and the following questions' coding procedure.

One of the critical questions asked following the main writing was whether the writing prompt was difficult to understand or not. The statements reported by the participants coded at three levels (almost no difficulty, somewhat difficulty, and significant difficulty). The frequency analysis revealed that understanding the writing prompt was not difficult for the participants (Table 18).

Table 18. The Perceived Difficulty of the Prompt across the Groups

	Experimental	Control
(Almost) No Difficulty	56	53
Somewhat Difficult	6	12
(Significant) Difficulty	4	1

As seen in the preceding table, only five participants in total reported difficulty in understanding the prompt, and four were from the experimental group. It might be essential to understand the data further to understand whether this affected their writing performance. For this purpose, I checked the following retrospective question (see the next paragraph) to see whether these participants had difficulty in writing, as well. Only Participant 307's answer was coded as significant difficulty for the main writing. She

reported that the words were difficult for her, and her writing was repetitive. However, when I checked her writing, I saw that she successfully completed the writing task; for example, she said, "the outdoor sports club and theatre club should be chosen. With the remaining five thousand, small needs can be purchased for the music club...". Moreover, it would be quite acceptable to encounter a few participants who had difficulty understanding and answering the prompt. Hence, I decided to keep her data.

The other question I asked was whether the participants had difficulty in writing the main passage. Again, the responses were coded at three (almost no difficulty, somewhat difficulty, and (significant) difficulty) levels. The frequency analysis revealed that most of the participants did not perceive the writing activity in the study as challenging. However, the experimental group might have experienced less difficulty than the control group. For example, less participants reported significant ($n = 16$ versus $n = 25$), but more participants reported almost no difficulty ($n = 21$ versus $n = 16$) in the experimental group (Table 19). This may be a tentative indication of the benefit of the writing task.

Table 19. The Perceived Difficulty of Main Writing across the Groups

	Experimental	Control
(Almost) No Difficulty	21	16
Somewhat Difficult	29	26
(Significant) Difficulty	16	25

I also checked whether the difficulty experienced during the main writing was associated with the difficulty in general writing. For this purpose, I excluded the other

and irrelevant answer categories from the general writing coding. And then, I took the percentage of the frequency values both in general writing (see Table 13 above) and in main writing (see Table 19). The comparisons revealed that the writing tasks in the study does not seem to be perceived as more difficult (Table 20). Actually, it seems that the main writing was perceived easier. For example, 27.82 % of the participants perceived no difficulty in main writing, but only 14.41 % perceived “no difficulty” in general writing. Still, this was not reflected in inferential analysis. Although it was not very strong, the Spearman correlation revealed a significant association between reporting difficulty in general writing and main experimental writing, $r(107) = .32, p = .001$. Hence, although the percentage of those who reported difficulty in general writing was higher, overall data indicated that the difficulty level in general and main writing was associated.

Table 20. The Percentage of Difficulty in General Writing Compared to the Difficulty in Main Writing

The Level of Difficulty	General Writing	Main Writing
(Almost) No Difficulty	14.41 %	27.82 %
Somewhat Difficult	52.25 %	41.35 %
(Significant) Difficulty	33.33 %	30.83 %

Another important thing to consider might be whether the participants who started the preparatory school at different proficiency levels experienced different difficulty levels while writing the main experimental passage. The investigation of the data with this respect revealed that the distributions were similar across the initial starting levels. For example, four out of twelve participants who started at relatively

higher levels (Track 3 and Track 4) reported having almost no difficulty in writing. Still, as I said above, as part of the language background control, I will check the starting proficiency level of the participants during the main analysis.

Turning back to retrospective reflection responses, the perceived difficulty during main writing was further analyzed to get insight into the source of the difficulty. The investigation revealed that the experimental group had more difficulty finding a solution for allocating the money, but the frequency was not so high in both groups (Table 21).

Table 21. The Reported Sources of Difficulty in Main Writing

Difficulty in Writing	Experimental	Control
No Difficulty	21	16
General English	7	3
Grammar	1	1
Vocabulary	6	6
General Writing Proficiency	4	5
Content/Background Gap	9	12
Content Organization	-	4
The Number of Words	5	9
Duration	-	-
Taking Perspective	3	1
Finding a Solution	9	2
Other	1	8

The other reflective question investigated whether the participants found an opportunity or preferred reviewing their passages. As seen in Table 22, most of the participants preferred reviewing their passages, and the review pattern of the passages was similar to each other.

Table 22. The Review Pattern of the Groups

Reviewing After Writing	Experimental	Control
Yes, I did	47	51
No, I did not prefer	12	10
No, there was a constraint	3	4
Other / Not Clear	4	2

After checking this frequency pattern, I also considered if using a computer rather than a paper would make a difference with respect to review/revision processes. For this purpose, I chose four students, two from the experimental and two from the control groups, considering the similarity in their EF performance. Participant 566 (EF z -score = 1.08) and Participant 403 (EF z -score = - 1.07) were from the control group, but Participant 625 (EF z -score = .085) and Participant 911 (EF z -score = -1.29) were from the experimental group. I shared the pictures of the passages they wrote, reminded the writing prompt, and asked them to respond to three questions via an online form. The English translation of the first two questions were:

Did using a pen-paper while writing this text have an impact on your reviewing and editing? What difference would it make if you typed the same text on the computer, especially in terms of editing the text you wrote?

The answers to this question revealed that three participants would prefer using pen-paper rather than writing it on a computer. The main reason for Participant 403 was using pen-paper allowed her to translate her ideas more comfortably, for Participant 911 was looking at the screen for a long time might make her tired, and for Participant 566, using pen-paper made her feel psychologically more comfortable and helped her focus on the topic, and writing and editing in the computer would be difficult for (and distract) her. However, Participant 625 reported that she writes faster by using a computer, and it would be better because tracking and editing the text (errors) would be easier and she would have more time for reviewing.

In a follow-up question, thinking that writing by using a paper can influence their evaluation processes before transcribing, I focused on idea generation and translation (also called formulation) before writing (see Hayes, 1981). The English translation of the question was:

Did writing the text using pen and paper cause you to think more before putting your thoughts down? How?

The answers to this question revealed that using pen-paper seemed to encourage the participants to think more. Participant 403 suggested that rather than thinking more, using pen-paper allowed her to plan by taking short notes and not to forget her ideas. Participant 911 similarly said, using paper, via taking short notes, allowed her to think more about the topic. Participant 566 reported that using paper helped her to write by thinking in more detail. Regarding Participant 625, who had said that using a computer could be more beneficial for her, she shared that with respect to thinking, using a paper did not matter much, and she would allocate around the same duration for thinking if she

used a computer. Based on these answers (and verbal conversations with many of my current students in the same institution and at the same proficiency level), I tend to think that writing by using pen and paper does not, at least, create significant constraints for those students.

Turning back to the retrospective questionnaire, the last reflection response I collected was about participant thoughts about how their passages could be evaluated. The most observable difference was that more participants in the control group ($n = 34$) than the experimental group ($n = 25$) thought that their passages would be evaluated in the classical testing context (Table 23). This may indicate that the experimental task created a relatively more authentic context. Additionally, although it was not one of the main aims of this study, I considered 41 out of 133 participants who found the main writing significantly difficult (see Table 19 above) and explored whether those 41 participants were mostly among the participants who expected that the evaluation would be done by an English Writing (Teacher). Cross-tabulation analyses revealed that this was the case for 22 out of those 41 participants.

Table 23. The Predictions Made about the Evaluation of Main Writing

Evaluation for	Experimental	Control
Perspective Taking	7	2
Ideas	16	13
English Writing (Teacher)	25	34
Other/Not Clear	14	16

6.1.3 Manipulation check analyses

I checked the efficiency of the manipulation in four different ways. Two were based on self-reports collected via the retrospective questionnaire (a five-point scale and an open-ended question). The other two were the coded PT and brainstorming activity notes, and the employment of EM and SM markers (see Manipulation Check Measurements part for further details) in the main passages. I will shortly remind the main experimental task and then report the analyses for the manipulation check in the above order.

To remind that, the experiment's main manipulation was directing the novice writers' attention to the audience of their passages. The participants in the experimental group were exposed to a writing task content that emphasized the audience's existence (e.g., names were mentioned in the vignettes), They were encouraged to take the audience's perspective (Theatre, Outdoor Sports, and Music Club members) and write whatever came to their minds in 3 minutes. On the other hand, the participants in the control group were exposed to relatively more neutral content (e.g., no names were mentioned in the vignettes) and were not encouraged to take the audience's perspective. They were just asked to write whatever came to their minds in 3 minutes. Having completed this activity and processed the prompt, all the participants were asked to write an around 250 words opinion passage.

Following this, the participants reflected on their PT experience. They considered a five-point scale and reported whether they tried taking the club members' perspective (1 = *not at all*, 5 = *completely so*) and whether they found it difficult (1 = *not (difficult) at all*, 5 = *completely so*) when they read the messages (Appendix F). With respect to finding it difficult, the comparison of the groups revealed that although

the overall mean value of the experimental group was lower ($M = 1.75$, $SD = .774$) than the control group ($M = 1.93$, $SD = .804$), both groups were statistically similar to each other, $t(131) = 1.263$, $p = .209$ (hence, it was not so difficult for both groups). However, the experimental group tried to take the perspective of club members and to understand their expectations significantly more ($M = 4.323$, $SD = .664$) than the control group ($M = 3.681$, $SD = .824$), $t(131) = 4.942$, $p < .001$. In other words, although it was equally difficult/easy for all participants, the manipulation successfully encouraged the experimental group to try to take the audience's perspective.

Additionally, in the retrospective questionnaire, I asked the participants to state whether they took their audience's perspective and whether doing this was influential in their writings. Again, the answers were coded for frequency comparisons (please see the Retrospective Questionnaire Coding section above and Appendix K below for the coding procedure). The data revealed that much more participants took an audience perspective in the experimental group ($n = 61$) compared to the control group ($n = 38$) (Table 24). To illustrate, Participant 582 from the control group reported that “Onları pek dikkate almadım çünkü verilen para belli imkan belli.”, but the Participant 752 from the experimental group reported that “Evet dikkate aldım. Örneğin bütçeyi ayarlarken hem rektör hem kulüp üyeleri hem de kulak öğrencilerin tarafından düşünerek yazmaya çalıştım.”

Table 24. The Distribution of the Reported Audience PT Across the Groups

	Experimental	Control
No Perspective Taking	4	25
Overall Reader Perspective	36	17
Rector Perspective	4	4
Club Members Perspective	11	10
Both Rector and Club Members	10	7
Other/Not Clear	1	1

To get further insight into the influence of PT manipulation, I also coded the responses for the second part of this question (whether PT influenced their writings). This coding also revealed that the influence of PT was felt more in the experimental ($n = 49$) than in the control group ($n = 27$) (Table 25).

Table 25. The Reported Influence of Audience PT on Writing across the Groups

	Experimental	Control
No influence of PT	17	35
The Influence of PT on Problem Solution	20	18
The Influence of PT on Text-Based Processes	11	2
The Influence of PT on Non-Specified Processes	18	7
The Influence of PT with Accompanied Difficulty	-	2

Thirdly, to get more concrete evidence, I checked the manipulation by considering the 3 minutes reaction of the participants to the presented vignettes (PT in the experimental and brainstorming in the control condition). I created a guideline (Appendix M) and coded these 3 minutes (9 minutes for three clubs in total) of PT or brainstorming notes written down by the participants. The results were obviously in accordance with the self-reports. Most of the participants from the experimental group (48 out of 66) tried to take the perspective of all (outdoor, music, and theatre) club members by imagining their situation or by directly putting themselves into the shoes of them. On the other hand, only 1 participant from the control group did this for all three clubs. Moreover, when the strength of the coded state PT statements (even if it was done for a single club) was evaluated on a 7-point scale, the strength of the experimental group's state PT was much higher ($M = 11.86$, $SD = 5.59$) than the control group ($M = 3.21$, $SD = 3.41$), $t(131) = 10.800$, $p < .001$.

The last manipulation check investigated the consideration of audience while writing. The question was whether the experimental task differentially influenced the frequency of the employment of EM and SM markers in the passages. Because the distribution of EM and SM markers were negatively skewed, which highly violated the normality assumption of parametric tests, non-parametric rather than parametric tests were conducted for the analyses of the relative EM and SM marker employment.

Before starting this last manipulation check analyses, I considered a concern about the employment of EM and SM: The employment of these markers may be influenced by the department/discipline of the writers. Hence, although our participants were from the EFL preparatory school and did not start to take their departmental

lessons, I wanted to check if their intellectual tendencies may be influential. For this purpose, I conducted a Kruskal-Wallis test and compared Mathematics/Physical Sciences, Verbal/Social Sciences, and Equal Weight students with respect to the employment of all types of EMs and SM markers. None of the comparisons revealed a significant difference across the groups (All P s > .1).

After eliminating this possibility, I compared the normalized frequency⁵⁰ of EMs and SM⁵¹ markers across the experimental and control groups (Table 26) by conducting nonparametric tests. The aim was to answer whether directing the novice EFL writers' attention to the audience increases the employment of explicit EMs and decreases the employment of SM markers. To start with, the mean values of almost all EM subcomponents (except from Appeal to Audience EM and Second Pronoun EM for General Audience) were higher, but SM subcomponents were lower in the experimental group. Additionally, Mann-Whitney U test results revealed significant differences between the experimental and control conditions. SMs were employed more by the control (*Mean Rank*: 76.68) than the experimental group (*Mean Rank*: 57.17), $U = 1562$, $Z = -2.919$, $p = 0.004$, but EMs were employed more by the experimental (*Mean Rank*: 73.7) than the control group (*Mean Rank*: 60.4), $U = 1768$, $Z = -2.001$, $p = 0.045$. The most important source of the difference in the EM use was its subcategory *Inclusive EM*

⁵⁰ As mentioned in the methodology section, I used T-Unit normalized EM and SM data in this study. Some researchers prefer reporting the data normalized over 1000 words, etc. I also created these kinds of alternative data (e.g., normalized per 1000 words, the data which considered the repeated EMs with T-Units by excluding the identical ones, the data which gives higher weight to some EMs such as directives (Midgette et al., 2008), the data which prioritizes inclusive “we” over “second pronoun you”, etc.) and checked if they create different results. Seeing that they also produced similar results, I decided to stick to my initial decision.

⁵¹ In these inferential statistical tests, if not otherwise stated, SM and EM refer to the total frequency (rather than the subcomponents) of these markers.

(e.g., we, us, our) which was employed more by the experimental (*Mean Rank*: 74.11) compared to the control group (*Mean Rank*: 60), $U = 1742$, $Z = -2.166$, $p = 0.03$. Thus, the experimental task seems to increase the employment of EM (particularly Inclusive EMs) but decrease SM.

Table 26. The Standardized Frequency of EMs and SM Markers Across the Experimental and Control Groups (Total frequency values for each subcategory are given in parentheses)

	Experimental	Control
Inclusive EM (Rector)	8.23 (138)	4.12 (64)
Inclusive EM (Club Members)	0.51 (8)	0.29 (5)
Inclusive EM (Both Rector and Club Members)	1.22 (17)	0.64 (13)
Inclusive EM (General Audience)	2.62 (40)	1.32 (22)
Total Inclusive EM*	12.59 (203)	6.38 (104)
Second Pronoun EM (Rector)	0.38 (5)	0.26 (4)
Second Pronoun EM (Club Members) *	1.58 (28)	0.06 (1)
Second Pronoun EM (Rector and Club Members)	-	-
Second Pronoun EM (General Audience)	0.51 (7)	1.76 (33)
Total Second Pronoun EM	2.48 (40)	2.09 (38)
Direct EM	3.25 (51)	2.78 (41)
Personal Aside EM	-	-
Question EM	0.32 (4)	0.29 (5)
Appeal to Audience EM	0.51 (7)	0.76 (12)
Total (Direct, Personal Aside, Question and Appeal)	4.09 (62)	3.83 (58)
Total EM*	19.15 (305)	12.29 (200)
Typical SM (I, me, my) *	16.84 (254)	27.49 (372)
Exclusive SM	0.05 (1)	1.66 (33)
Total SM*	16.90 (255)	29.15 (405)

*The *Mann-Whitney U* revealed a significant difference between the groups ($p < .05$).

On the basis of these analyses, I suggest that the experimental manipulation, which aimed to direct the participants' attention to the audience in the experimental groups, was successful. Moreover, the significantly less employment of SM but more employment of EM by the experimental group may indicate that the audience processes were more active while writing in the experimental group.

Now that I prepared the data, checked the background similarity across the groups, investigated their reaction to the task, and become sure that the manipulation worked well, I will focus on the main questions of the study and try to understand the influence of the experimental audience-oriented task on the writing patterns of these novice EFL writers in detail.

6.2 Main analyses

To remind that, the primary aim of this study was to create an experimental task/condition that directs the attention of novice Turkish EFL writers to the audience and to investigate its effect by considering EF capability and trait PT tendency of the participants. For this purpose, I designed an audience-oriented experimental and a non-audience-oriented control task. The tasks included four components: vignettes, pictures, a kind of short brainstorming (control) or PT (experimental) section, and a prompt (see the Materials section). For example, in the audience-oriented experimental condition, after reading each vignette/message, the participants were asked to take the perspective of the club members (think about their expectations, thoughts, feelings, etc.) and write whatever came to their minds within 3 minutes. However, in the non-audience-oriented control condition, after reading each vignette/message, the participants were asked to

simply write whatever came to their minds within 3 minutes. It was obvious that asking to take the audience's perspective would increase the audience orientation of the task and this was confirmed by four different manipulation check analyses (see Manipulation Check Analyses part above).

Before starting the main analyses, I considered the possible influence of the department/discipline of the EFL writers and checked whether their intellectual tendencies may interact with the experimental or control tasks and reveal different patterns. For this purpose, I conducted two univariate ANOVA analyses, one for the overall writing quality and one for the level of persuasiveness. The fixed variables were the main study areas (Mathematics/Physical Sciences, Verbal/Social Sciences, and Equal Weight) and the task conditions (experimental and control). The results did not reveal any significant interactions (All P s > .1).

Now that this concern has been eliminated within this study's context, I will focus on the main analyses. I will start with the effects of task type (audience-oriented experimental versus non-audience-oriented control) and EF capability (high versus low) on the overall writing quality and the level of persuasiveness of the passages. Then, I will focus on whether trait PT (high versus low) moderates the effect of the task type (audience-oriented experimental versus non-audience-oriented control), and finally, if this is the case, I will consider whether trait PT can compensate for the lack of sufficient EF resources when a novice low EF capability EFL writer was exposed to an audience-oriented experimental task.

6.2.1 The effects of task and EF capability on novice EFL writers' performance

In this part, I will consider the effects of task type (audience-oriented experimental versus non-audience-oriented control) and EF capability (high versus low) on the overall writing quality and the level of persuasiveness of the passages written by novice EFL writers and test the first and the second hypotheses of this study (the details are given below).

6.2.1.1 The effects of the task and EF on overall writing quality

The study's first hypothesis predicted that the experimental audience-oriented writing task would increase the overall writing quality performance of novice EFL writers with high EF capability.

To test this hypothesis, I conducted a 2-way factorial ANOVA. The dependent variable was the overall writing quality scores of the participants. The independent variables were EF capability (low versus high) and the task (audience-oriented experimental versus non-audience-oriented control). Most important for our present concern, the results revealed an interaction between the EF (high versus low) and task (experimental versus control) conditions, $F(1, 116) = 10.083, p = .002$, (Fig. 12). Also, there was a main effect of EF on the overall writing quality. The group with relatively higher EF ($M = 71.45, SD = 7.42$) performed better than the group with relatively lower EF ($M = 66.09, SD = 8.71$), $F(1, 116) = 14.182, p < .001$. However, there was no main effect of the task. The experimental ($M = 69.23, SD = 8.82$) and the control groups ($M = 68.33, SD = 8.21$) performed similarly, $F(1, 116) = .096, p = .757$.

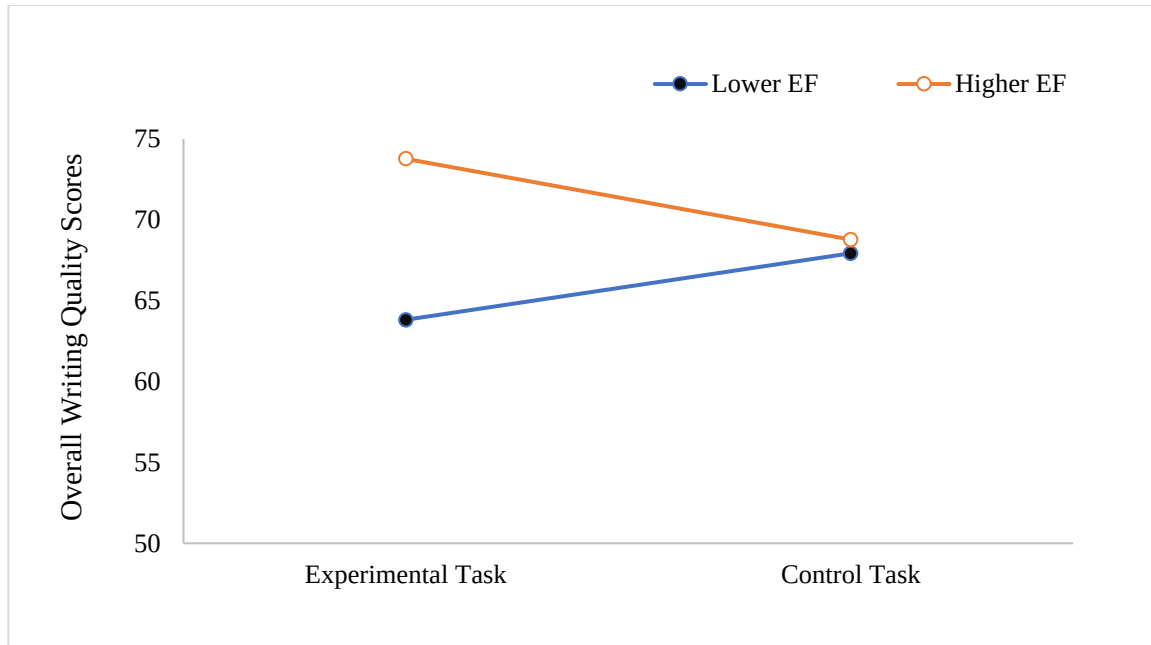


Fig. 12 The overall writing quality performance of relatively higher and lower EF groups in the experimental and control conditions

The following between subject t-test analyses, which aimed to locate the source of the interaction, revealed a differential influence of the audience-oriented experimental task on the participants with relatively low EF ($M = 63.83$, $SD = 8.26$) and relatively high EF ($M = 73.78$, $SD = 6.45$) capabilities, $t(57) = 5.195$, $p < .001$. The experimental task was much more beneficial for improving the overall writing quality of the passages written by the higher EF capability group. When it comes to the control task, it did not reveal a significant difference between the low EF ($M = 67.94$, $SD = 8.75$) and the high EF ($M = 68.79$, $SD = 7.66$) groups, $t(59) = .398$, $p = .692$. Both EF groups performed similarly.

In addition to this, I also split and compared the performance of low EF and high EF groups in the experimental versus control conditions. The between-subject t-test results revealed that the high EF participants in the experimental condition ($M = 73.78$,

$SD = 6.45$) performed better than the high EF participants in the control condition ($M = 68.79$, $SD = 7.66$), $t(58) = 2.742$, $p = .008$. When it comes to the low EF groups, the overall writing quality performance of those in the experimental condition ($M = 63.83$, $SD = 8.26$) was not significantly different from the performance of those in the control condition ($M = 67.94$, $SD = 8.75$), $t(58) = 1.855$, $p = .069$.

Hence, the crucial source of the interaction was the better performance of the participants with higher EF capability in the audience-oriented experimental condition. This group performed better than the low EF group in the experimental condition and the high EF group in the control condition.

Following this observation, I split the data and focused only on this high EF group. I conducted further analyses to investigate whether the task (audience-oriented experimental versus non-audience -oriented control) can still predict the overall writing quality after fixing/controlling their language background and trait PT tendencies; doing this would add another dimension to the previous analyses.

For this purpose, I conducted a hierarchical multiple regression analysis. The model consisted of four levels. In the first level, the initial proficiency exam scores that the participants took at the beginning of the academic year, and in the second level, the current vocabulary, grammar, and reading comprehension scores were added. Then, the participants' trait PT data was included in the third level. Finally, the dummy coded audience-oriented experimental versus non-audience-oriented control task conditions were added.

The results revealed that the model⁵² was significant at first, $F(1, 56) = 6.403$, $p = .014$, $R^2 = .103$, at second, $F(4, 53) = 4.279$, $p = 0.004$, $R^2 = .244$, at third, $F(5, 52) = 4.039$, $p = .004$, $R^2 = .280$, and at fourth levels, $F(6, 51) = 4.958$, $p < .001$, $R^2 = .368$. The model explained around 37% of the variance for this higher EF group and 8.8 % of this variance was added by the task (Table 27).

The further investigation indicated that the initial (starting) proficiency score contributed to the model at the first, $p = .014$ ($b = .32$)⁵³, at the second, $p = .043$ ($b = .266$), at the third, $p = .042$ ($b = .263$), and at the fourth levels, $p = .044$ ($b = .247$). The vocabulary, grammar and PT performance did not contribute to the model at any levels (All $P_s > .1$). Reading performance of the participants contributed to the model at the second, $p = .014$ ($b = .311$), at the third, $p = .008$ ($b = .339$), and at the fourth levels, $p = .007$ ($b = .322$). Most critically, the experimental condition predicted the overall writing quality scores after fixing/controlling the language background and trait PT of the participants, $p = .01$ ($b = .299$, *unstandardized beta* = 4.48)⁵⁴. Hence, the model indicated that if the experimental task is presented to a participant with high EF, his/her

⁵² To note, the investigation of the data revealed that the basic assumptions of linear regression were met. First of all, the dependent variable (writing scores) was normally distributed (Kolmogorov-Smirnova, $P_s > .1$). P-Plots also revealed that the data were normal. Secondly, the Maximum Cook's distance value was below 1, so no outlier was detected. Third, no multicollinearity was detected among the predictors (all $P_s < .7$) and VIFs were below 10, and tolerance values were greater than .1. And fourth, the plot of standardized residual values indicated that the data were homoscedastic (almost no heteroscedasticity), and standard residual values were within -3 +3 range. Thus, the relationship between the predictor and outcome variables was linear. I will not report the assumption check for the following regression analysis if there is no exceptional case.

⁵³ I will report the standardized rather than unstandardized/raw beta values, if not otherwise stated.

⁵⁴ I conducted the same analysis for low EF group to explore the data further, but the results did not reveal any significant prediction of the task.

overall writing quality score may be increased by around 4.5/100 points (after the language background and trait PT tendency were fixed).

Table 27. Overall Writing Quality Performance of Higher EF Participants Predicted by Experimental Task Condition Controlling the Language and Trait PT Background

	Model 1	Model 2	Model 3	Model 4
Initial Proficiency	$b = 0.32^*$	$b = 0.266^*$	$b = 0.263^*$	$b = 0.247^*$
Grammar		$b = 0.112$	$b = 0.223$	$b = 0.160$
Vocabulary		$b = 0.169$	$b = 0.192$	$b = 0.263$
Reading		$b = 0.311^*$	$b = 0.339^{**}$	$b = 0.322^{**}$
Trait PT			$b = 0.197$	$b = 0.181$
Task Condition				$b = 0.299^{**}$
The Variance Explained R^2	10.3 %*	24.4 %**	28 %**	36.8 %***
ΔR^2 Change	.103	.141	.036	.088

* $p < .05$. ** $p < .01$. *** $p < .001$. b = standardized beta values

Overall, the analyses confirmed the first hypothesis of this study. The experimental audience-oriented task directing the attention of novice EFL writers to the audience can improve the overall writing quality performance of the participants with a relatively higher EF capability. Moreover, the contribution of the task, which added 8.8 % to the total variance, is beyond the L2 background and trait PT tendencies of those novice Turkish EFL writers.

6.2.1.2 The effects of the task and EF on the level of persuasiveness

The study's second hypothesis predicted that the experimental audience-oriented writing task would increase the level of persuasiveness performance of novice EFL writers with high EF capability.

Again, I conducted a 2-way factorial ANOVA to test this hypothesis. Most important for our present concern, there was a marginally significant interaction between EF capability (high versus low) and task (audience-oriented experimental versus non-audience oriented control) conditions⁵⁵, $F(1, 116) = 3.360$, $p = .069$, (Fig. 13). Other than that, the results revealed no main effect of EF on the level of persuasiveness, $F(1, 116) = 1.182$, $p = .279$. The performance of the high EF group ($M = 3.43$, $SD = 1.21$) was not significantly higher than the low EF group ($M = 3.20$, $SD = 1.11$). The main effect of the task was not significant, as well. The experimental ($M = 3.39$, $SD = 1.29$) and the control groups ($M = 3.25$, $SD = 1.03$) performed similarly, $F(1, 116) = .347$, $p = .557$.

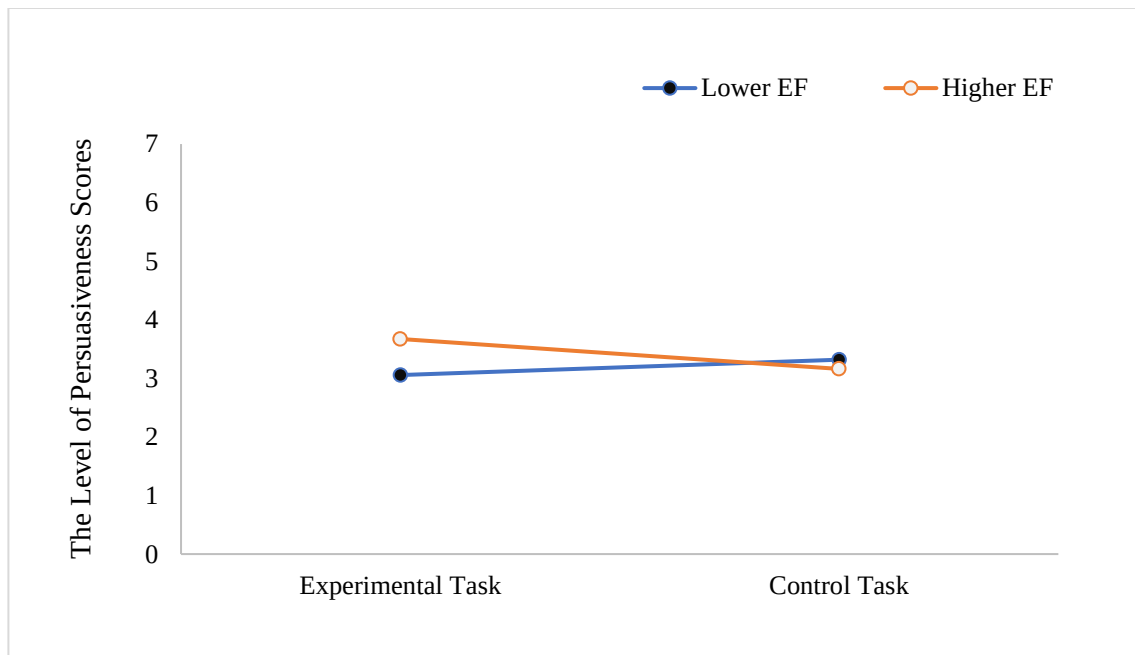


Fig. 13 The level of persuasiveness performance of relatively higher and lower EF groups in the experimental and control conditions

⁵⁵ I conducted the same ANOVA analyses with the subcomponents of this EF composite score (WM composite, AC composite, flanker, antisaccade, symmetry, and operational span scores) in order to check whether there was a significant interaction of the task with them. However, no significant results were revealed (All P s > .1).

The following two-tailed between-subject t-test analyses, which aimed to locate the source of the interaction, revealed a significant difference between low EF and high EF participants who were exposed to the audience-oriented experimental task. The level of persuasiveness performance of high EF capability group was better ($M = 3.67$, $SD = 1.35$) than the low EF group ($M = 3.06$, $SD = 1.15$), $t(57) = 1.867$, $p = .034$ (one-tailed)⁵⁶. However, the non-audience-oriented control task did not reveal such a significant difference between the low EF ($M = 3.32$, $SD = 1.08$) and the high EF groups ($M = 3.16$, $SD = .97$), $t(59) = .593$, $p = .555$. Both EF groups performed similarly.

In addition to these, I also split and compared the level of persuasiveness performance of low EF and high EF groups in the experimental versus control conditions. The between-subject t-test results revealed that the high EF group in the experimental condition ($M = 3.67$, $SD = 1.35$) performed similarly to the high EF group in the control condition ($M = 3.16$, $SD = .97$), $t(58) = 1.658$, $p = .103$. Lastly, although the mean level of persuasiveness values revealed a reversed pattern for the low EF group, their performance in the control condition ($M = 3.32$, $SD = 1.08$) was not significantly different from the low EF group in the experimental condition ($M = 3.06$, $SD = 1.15$), $t(58) = .911$, $p = .366$.

Overall, the second hypothesis of this study was confirmed at a marginal level. The experimental audience-oriented task directing the attention of novice EFL writers may improve the level of persuasiveness performance of the participants with a

⁵⁶ I considered one-tailed analysis because this was explicitly predicted by the second hypothesis.

relatively higher EF capability, but further investigation may be needed to make a clear conclusion⁵⁷.

6.2.2 The effects of task and trait PT on novice EFL writers' performance

Now, I will focus on the second question and consider whether trait PT (high versus low) moderates the effect of the task type (audience-oriented experimental versus non-audience-oriented control) on the overall writing quality and the level of persuasiveness of the passages written by novice EFL writers and try to test the third and fourth hypotheses of this study (the details are given below).

6.2.2.1 The effects of the task and trait PT on overall writing quality

The study's third hypothesis predicted that the experimental audience-oriented writing task would increase the overall writing quality performance of novice EFL writers with high trait PT tendency.

To test this hypothesis, I conducted a 2-way factorial ANOVA. The dependent variable was the overall writing quality scores of the participants. The independent variables were the trait PT (low versus high) and the task (audience-oriented experimental versus non-audience-oriented control). The results revealed that there was a statistically significant interaction between trait PT and the task, $F(1, 129) = 4.465, p =$

⁵⁷ Similar to overall writing quality, I further explored the high EF group's data by conducting a hierarchical regression analysis in which I controlled the language background and trait PT. The results revealed that the dummy-coded task conditions did not predict the level of persuasiveness after controlling these variables ($p > .1$). Thus, I will approach this marginally significant finding with precaution and consider the literature before concluding.

.037, (Fig. 14). Other than that, there was no main effect of trait PT, $F(1, 129) = 0.070$, $p = .792$, or the task on the overall writing quality, $F(1, 129) = 1.511$, $p = .221$.

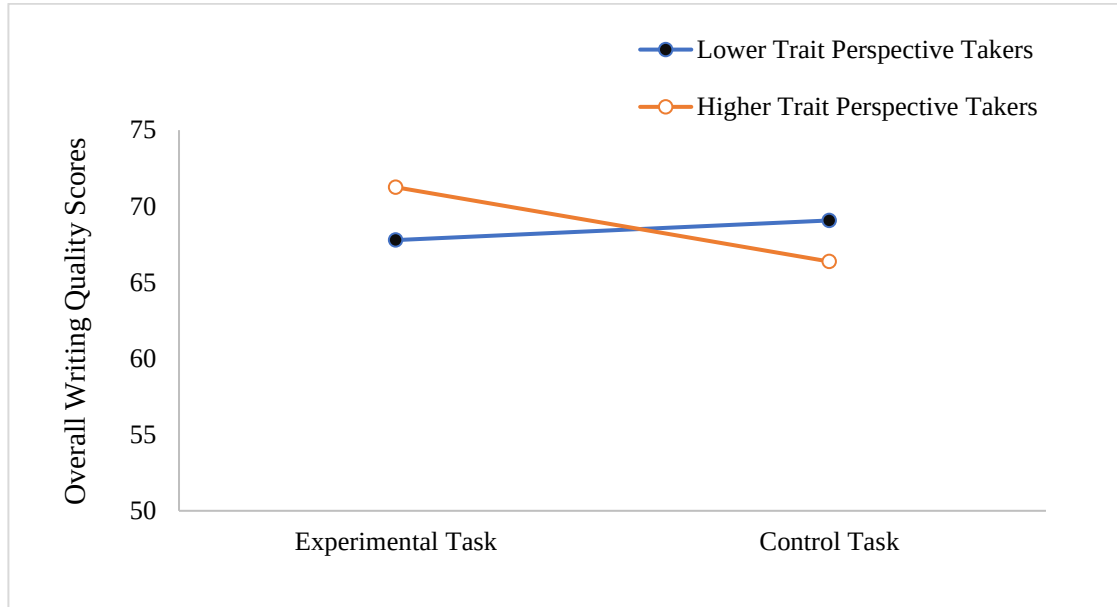


Fig. 14 The overall writing quality performance of relatively higher and lower PT groups in the experimental and control task conditions

The following between-subject t-test analyses revealed that the source of this interaction was the positive influence of the audience-oriented experimental task on the participants with high trait PT tendency. The overall writing quality of the high trait PT group in the experimental condition was improved more ($M = 71.25$, $SD = 7.14$) than of the high trait PT group in the control task condition ($M = 66.37$, $SD = 8.69$), $t(61) = 2.438$, $p = .018$. None of the other t-test comparisons were significant (All P s $> .1$). Hence, exposing the high trait PT participants to an audience-oriented experimental task, compared to a non-audience-oriented control task, can increase their overall writing quality performance more.

Following this observation, I focused on this high trait PT group and conducted further analyses to investigate whether the task (audience-oriented versus non-audience-oriented) can still predict the overall writing quality after fixing/controlling their language background and EF capabilities.

For this purpose, I again conducted four level hierarchical multiple regression analyses. In the first level, the initial proficiency exam scores that the participants took at the beginning of the academic year, and in the second level, the current vocabulary, grammar, and reading comprehension scores were added. Then, the participants' EF scores were included in the third level. Finally, the dummy coded experimental versus control conditions were added.

The model was not significant at first, $F(1, 52) = 2.651, p = .110, R^2 = .049$, but significant at second, $F(4, 49) = 4.581, p = .003, R^2 = .272$, at third, $F(5, 48) = 4.947, p = .001, R^2 = .34$, and at fourth levels, $F(6, 47) = 5.457, p < 0.001, R^2 = .411$ (Table 28). It explained 41 % of the variance in this higher trait PT group.

Further investigation of predictors revealed that the initial (starting) proficiency score did not contribute to the model at the first, $p = .11$ ($b = .22$) and at the second, $p = .063$ ($b = .238$), but contributed at the third, $p < .05$ ($b = .264$), and at the fourth levels, $p < .05$ ($b = .292$). The vocabulary performance of the participants did not significantly contribute to the prediction of main writing scores at the second or the third and contributed only at a marginal level at the fourth level, $p = .056$ ($b = .263$). The grammar performance did not contribute to the model at any levels (All $P_s > .1$). Reading performance of the participants contributed to the model at the second, $p = .005$ ($b = .366$), and at the third levels, $p = .01$ ($b = .324$), but its contribution was only

marginally significant at the fourth level, $p = .059$ ($b = .235$). EF was entered into the model at the third level, and it was shown that its contribution, after controlling the language-related background, was significant both at the third, $p < .05$ ($b = .288$) and at the fourth levels, $p < .05$ ($b = .291$). Most critically, the task conditions predicted the overall writing quality scores even after fixing/controlling the language background and EF of the participants, $p < .05$ ($b = .289$, *unstandardized beta* = 4.52). Hence, this model indicates that if the experimental task is presented to a high PT participant, his/her overall writing quality score may be increased by around 4.5/100 points⁵⁸.

Table 28. Overall Writing Quality Performance of the Higher Trait PT Group Predicted by Experimental Task Condition Controlling the Language and EF Background

	Model 1	Model 2	Model 3	Model 4
Initial Proficiency	$b = 0.22$	$b = 0.238$	$b = 0.264^*$	$b = 0.292^*$
Grammar		$b = 0.112$	$b = 0.223$	$b = 0.160$
Vocabulary		$b = 0.169$	$b = 0.192$	$b = 0.263$
Reading		$b = 0.366^{**}$	$b = 0.324^{**}$	$b = 0.235$
EF			$b = 0.288^*$	$b = 0.291^*$
Task Condition				$b = 0.289^*$
The Variance Explained R^2	4.9 %	27.2%***	34 %***	41 %***
ΔR^2 Change	.049	.223	.068	.071

* $p < .05$. ** $p < .01$. *** $p < .001$. b = standardized beta values

Overall, the analyses confirmed the third hypothesis of this study. The experimental audience-oriented task directing the attention of novice EFL writers to the audience can improve the overall writing quality performance of the participants with a relatively higher trait PT tendency. Moreover, the contribution of the task, which added

⁵⁸ To note, I conducted the same analysis for low trait PT group, but the results did not reveal any significant prediction of the task.

7.1 % to the total variance, is beyond the L2 background and EF capabilities of those novice Turkish EFL writers.

6.2.2.2 The effects of the task and trait PT on the level of persuasiveness

The study's fourth hypothesis predicted that the experimental audience-oriented writing task would increase the level of persuasiveness performance of novice EFL writers with high trait PT tendency.

To test this hypothesis, once again, I conducted a 2-way factorial ANOVA. The dependent variable was the level of persuasiveness of the participants. The independent variables were trait PT tendency (low versus high) and the task (audience-oriented experimental versus non-audience-oriented control). The results did not reveal a statistically significant interaction between trait PT and the task, $F(1, 129) = .895, p = .346$, (Fig. 15). Additionally, there was neither a main effect of the task, $F(1, 129) = 1.139, p = .288$, nor of trait PT. The level of persuasiveness performance of the relatively high trait PT ($M = 3.26, SD = 1.26$) and the low trait PT ($M = 3.44, SD = 1.14$) groups was similar to each other, $F(1, 129) = .782, p = .378$.

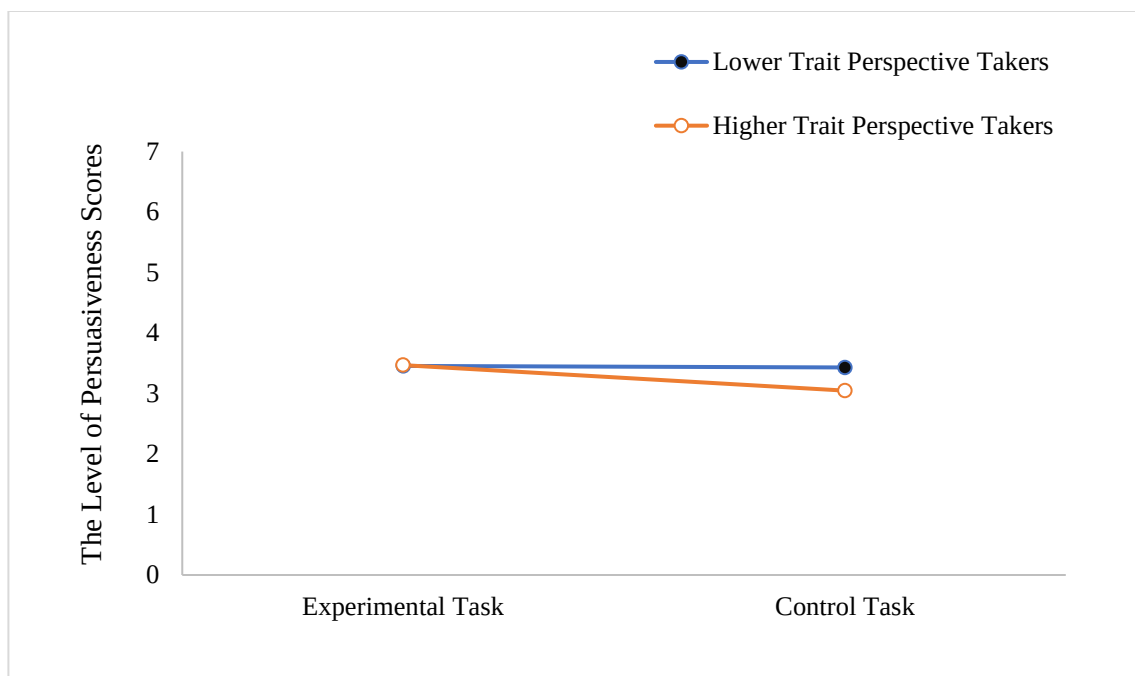


Fig. 15 The level of persuasiveness performance of relatively higher and lower PT groups in the experimental and control task conditions

Hence, the analyses rejected this study's fourth hypothesis. The experimental audience-oriented task directing the attention of novice EFL writers cannot improve the level of persuasiveness performance of the participants with a relatively higher trait PT ability.

6.2.3 The compensation of insufficient EF resources by Trait PT tendency

The analyses in the previous section indicated that trait PT can moderate the effect of the experimental audience-oriented writing task on the overall writing quality but not on the level of persuasiveness. I will now consider whether high trait PT tendency can compensate for the lack of sufficient EF resources by testing the fifth hypothesis and exploring the data further for the sixth hypothesis.

6.2.3.1 The compensation for increasing the overall writing quality

The previous analyses confirmed the involvement of trait PT in the experimental audience-oriented task processing with respect to its influence on the overall writing quality. Following this, the study's fifth hypothesis predicted that high trait PT tendency could compensate for insufficient EF resources (low EF) and increase the overall writing quality performance of novice EFL writers exposed to the experimental audience-oriented writing task.

This hypothesis was concerned with the participants in the audience-oriented experimental group with (low versus high) EF capability and trait PT tendencies. Hence, although I shared the descriptives of the control group, I focused on the experimental group and created four combinations from the participants with relatively high EF, low EF, high trait PT, and low trait PT. This resulted in low EF – low PT, low EF – high PT, high EF – low PT, and high EF – high PT groups. The most critical group which could help me test the fifth hypothesis was the low EF – high PT. If this group, whose EF capability was low but trait PT ability was high, performed similar to the high EF – low PT and high EF – high PT groups but performed better than the low EF – low PT group, it would indicate that high trait PT tendency can reduce the demands that emerged from the audience-oriented nature of the experimental task and can compensate for the lack of sufficient EF resources in this way.

The preliminary investigation of the data revealed that the mean overall writing quality of low EF – high PT participants in the experimental group was below the high EF – low PT and high EF – high PT but above the low EF – low PT participants (Table 29). Beyond these, the inferential univariate ANOVA analysis revealed a significant

main effect, $F(3, 55) = 10.787, p < .001$, and the Bonferroni corrected post hoc comparisons indicated that the source of this effect was the difference between high EF – low PT and low EF – low PT ($p = .001$), as well as between high EF – high PT and low EF – low PT groups ($p < .001$). However, the overall writing quality performance of low EF – high PT was not different from the other groups (all $P_s > .1$)⁵⁹.

Table 29. The Mean Overall Writing Quality Values as a Function of EF, trait PT, and Experimental Conditions

EF	PT	Condition	<i>N</i>	<i>Mean</i>	<i>SD</i>
Low EF	Low PT	Experimental	17	61.68	7.68
Low EF	Low PT	Control	22	68.00	8.70
Low EF	High PT	Experimental	10	67.50	8.26
Low EF	High PT	Control	11	67.82	9.26
High EF	Low PT	Experimental	13	74.39	6.81
High EF	Low PT	Control	11	71.14	7.62
High EF	High PT	Experimental	19	73.37	6.35
High EF	High PT	Control	17	67.27	7.52

Overall, the analyses seemed to partially confirm the fifth hypothesis of this study. Unlike the low EF – low PT group, whose overall writing quality performance was significantly lower than high EF groups, the low EF – high PT group's performance was similar to the high EF groups. However, although the low EF – high PT group's

⁵⁹ When I conducted the same post hoc test with LSD (Fisher's Least Significant Difference Test), the results revealed a significant difference between the low EF – low PT group ($p = .047$). Still, because many researchers did not recommend this test, I decided to stick with my initial decision and considered the Bonferroni corrected comparisons.

overall writing quality mean values were higher than the low EF – low PT group, it was not enough to statistically differentiate them. Thus, it seems that when novice Turkish EFL writers were exposed to an experimental audience-oriented indicated task, trait PT may only partially compensate for the insufficient EF resources.

6.2.3.2 The compensation for increasing the level of persuasiveness

Although it was dependent on the fourth hypothesis (trait PT contributes to the effect of the task on the level of persuasiveness), which was rejected by the analyses, I tested the sixth hypothesis to investigate the data with respect to whether high trait PT tendency could compensate for insufficient EF resources and increase the level of persuasiveness performance of novice EFL writers who were exposed to the experimental audience-oriented writing task. The aim was to explore the pattern and share the analyses.

For this purpose, similar to the overall writing quality data, I focused on the experimental group and created four different EF (high versus low) and trait PT (high versus low) combinations. The data revealed that the mean level of persuasiveness value of the low EF – high PT group was similar to low EF – low PT and both low EF groups' mean values seemed to be lower than the high EF groups' (Table 30). Still, univariate ANOVA analysis did not reveal a significant main effect, $F(3, 55) = 1.176, p = .327$.

Table 30. The Mean Level of Persuasiveness Values as a Function of EF, trait PT, and Experimental Conditions

EF	PT	Condition	<i>N</i>	<i>Mean</i>	<i>SD</i>
Low EF	Low PT	Experimental	17	3.09	1.20
Low EF	Low PT	Control	22	3.25	1.03
Low EF	High PT	Experimental	10	3.00	1.11
Low EF	High PT	Control	11	3.46	1.21
High EF	Low PT	Experimental	13	3.77	1.17
High EF	Low PT	Control	11	3.46	0.99
High EF	High PT	Experimental	19	3.61	1.50
High EF	High PT	Control	17	2.97	0.94

Hence, these exploratory analyses rejected this study's sixth hypothesis and confirmed that trait PT does not contribute to the level of persuasiveness within this study's context. More specifically, regarding the level of persuasiveness performance, when the participants with low EF – high PT were exposed to the experimental audience-oriented task, trait PT may not compensate for the lack of sufficient EF resources.

6.3 Additional analyses

I conducted a series of additional analyses by including the diagnostic writing of the participants. The first one focused on the overall writing quality of the diagnostic versus main experimental stage passages written by the participants whose trait PT and EF capabilities were relatively lower than the others (low EF – low PT group). Following

this, I checked the correlation between diagnostic writing performance, EF, and trait PT. Additionally, I checked the correlations between main writing performance and trait PT for the control and experimental groups separately.

6.3.1 Diagnostic versus main writing performance of low EF – low PT Group

The first additional analysis aimed to explore the overall writing quality performance of those EFL writers who were low both in trait PT and EF. As reported in the main analysis section (Table 29), the mean score of this group in the experimental condition ($M = 61.68$, $SD = 7.68$, $n = 17$) was lower than in the control condition ($M = 68.00$, $SD = 8.70$, $n = 22$). When I split this low EF – low PT group's diagnostic writing data in order to check whether this was related to their writing background (a possibility of unequal distribution), I saw that diagnostic writing performance of both those who were exposed to the control ($M = 67.50$, $SD = 8.62$) and those who were exposed to the experimental conditions ($M = 67.18$, $SD = 7.28$) were similar to each other and to the main writing performance of the same low EF – low PT control group ($M = 68.00$, $SD = 8.70$)⁶⁰. However, it seemed that the experimental condition, which directed their attention to the audience, reduced their overall writing quality performance.

I converted all diagnostic and main writing scores into z-values to investigate this further. The aim was to make these diagnostic and main overall writing quality scores statistically comparable and determine the performance of those participants relative to others. I then conducted a 2 (diagnostic versus main writing) by 2

⁶⁰ It may be important to note that I also checked the efficiency of the manipulation for this group. The investigation of self-reports and the frequency of EM and SM markers revealed that this group's audience processing pattern was similar to the other experimental groups (e.g., high EF – high PT). Thus, as the other experimental groups, this group was actively engaged with the audience during writing.

(experimental versus control) mixed ANOVA analysis. The within-subject variable was the z-transformation of diagnostic versus main overall writing quality scores. The results revealed an interaction between diagnostic versus main overall writing quality and the experimental versus control tasks, $F(1, 37) = 16.736, p < .001$. The within-subject t-test comparisons revealed that the source of this effect was this (low PT – low EF) group's performance decrease in the overall writing quality from the diagnostic ($mean\ z = -.40, SD = .86$) to the main writing ($mean\ z = -.82, SD = .90$) when they were exposed to the experimental condition, $t(16) = 3.346, p = .004$; but performance increase in the overall writing quality from the diagnostic ($mean\ z = -.36, SD = 1.02$) to main writing ($mean\ z = -.075, SD = 1.03$) when they were exposed to the control condition, $t(21) = -2.468, p = .022$ (Fig. 16). Furthermore, when I subtracted the diagnostic overall writing quality z-values from the main overall writing quality z-values (thus, a high z value would mean better main writing performance) of this low EF – low PT group and compared these mean z value differences across the experimental ($mean\ z\ difference = -.42, SD = .51$) and control ($mean\ z\ difference = .29, SD = .55$) conditions, a between-subject t-test analysis revealed a significant difference, $t(33) = -4.091, p < .001$. Thus, the performance of this group was improved by the control but worsened by the experimental task⁶¹. To note, the data did not reveal any significant results when I conducted the same series of these analyses for low EF – high PT groups (all P s > .1), hence, although insufficient, this may be another tentative support for the suggestion

⁶¹ One-sample t-test analyses confirmed this. The low EF – low PT control group's performance difference was above 0 ($t(21) = -2.468, p = .022$), but the low EF – low PT experimental group's performance difference was below 0 ($t(16) = -3.346, p = .004$).

that having high trait PT may partially compensate for the insufficient EF resources when an audience-oriented task overloads the system.

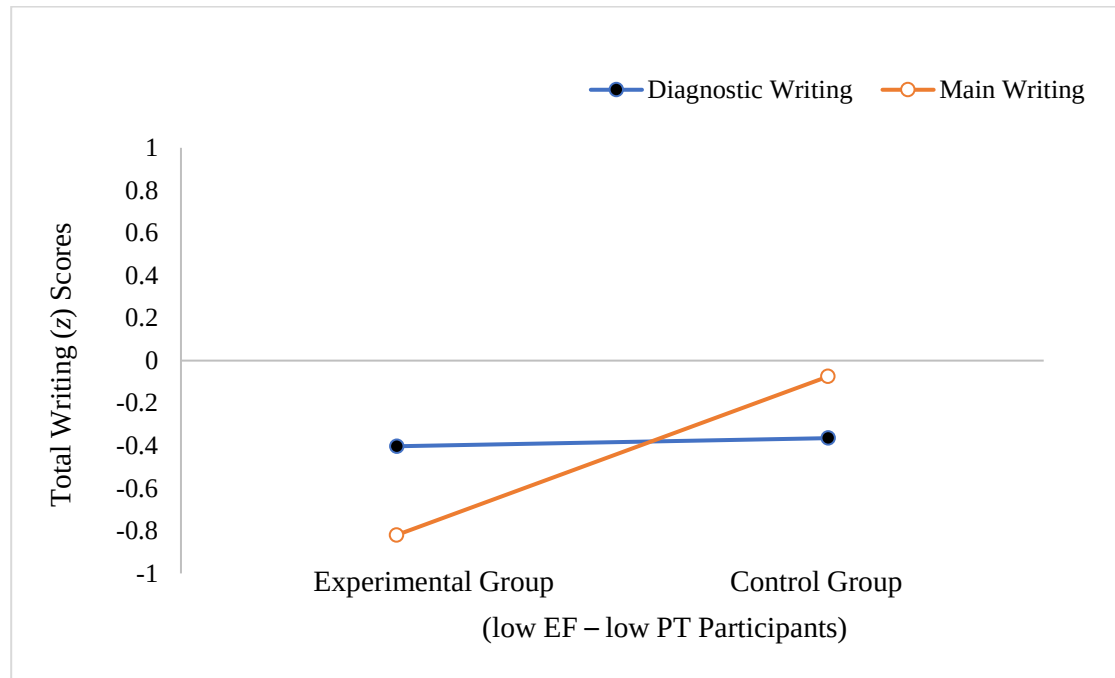


Fig. 16 The overall writing quality pattern revealed by relatively low PT – low EF group across their within subject diagnostic versus main overall writing quality and between subject experimental versus control conditions

6.3.2 Associations between diagnostic overall writing quality and trait PT and EF

I also conducted one-tailed Pearson correlations with pairwise exclusion of the missing data to see if there were relationships between diagnostic overall writing quality performance and the critical variables of this study, EF⁶² and trait PT. The aim was to check whether the data, in the absence of experimental manipulation, was in accordance with the literature. The results indicated that EF was correlated both with diagnostic

⁶² I also considered the correlations between EF and fundamental language background (vocabulary and grammar exam scores). The analyses revealed no significant associations (All P s > .1). This does not tell much because these tests assess LTM rather than online EF processes during grammar and vocabulary processing.

overall writing quality, $r(120) = .353, p < .001$, and trait PT, $r(120) = .191, p = .018$.

However, trait PT did not correlate with diagnostic overall writing quality, $r(133) = .106, p = .111$, (Table 31).

Table 31. The Associations Between Three Critical Variables of the Study: Diagnostic Overall Writing Quality, EF, and Trait PT

		EF	Trait PT
Trait PT	<i>r</i>	.191*	
	<i>sig.</i>	.018	
	<i>N</i>	120	
Diagnostic Overall Writing Quality	<i>r</i>	.353**	.106
	<i>sig.</i>	.000	.111
	<i>N</i>	120	133

* Correlation is significant at the 0.05 level (1-tailed).

** Correlation is significant at the 0.01 level (1-tailed).

Hence, although the data confirmed the association between the EF and writing performance, it did not reveal a possible association between the tendency to take the other's perspective (trait PT) and writing performance. Still, it may be important to consider that the participants of this study were novice EFL writers who were still struggling with basic language processes and who might be unable to employ higher-level processes such as audience PT while writing under normal circumstances. The potential influence of trait PT may not be activated because of this. In other words, they might need their attention to be directed to the audience via a task in order to employ their trait PT abilities, take the audience's perspective and write better passages.

Although I investigated this suggestion in the preceding main analyses section and revealed a causal effect of such a task, thinking of that this could add another

dimension, I split the control and the experimental groups' data to separately check the associations between trait PT and the main experimental overall writing quality.

Revealing no association between the trait PT and the overall writing quality, $r(67) = -.077$, $p = .267$, one-tailed correlational data of the control group replicated the above-mentioned observation. In contrast, the correlational data of the experimental group revealed an association between trait PT and the overall writing quality, $r(66) = .261$, $p = .017$. Moreover, one-tailed Fisher's z-test indicated that the correlation coefficient of the control group ($-.077$) was significantly different/stronger than the correlation coefficient of the experimental group ($.261$), $z = 1.94$, $p = .026$. Thus, creating an audience-oriented task and directing the writer's attention to the audience may have the potential to activate the trait PT and associate it with the overall writing quality.

CHAPTER 7

DISCUSSION

This study adopted a cognitive approach and investigated the efficiency of a task that directed the attention of novice upper-intermediate level Turkish EFL writers to their audiences. This was done by creating an audience-oriented experimental task and a non-audience-oriented control task. The most distinctive feature of these tasks was that the participants in the experimental group were asked to read short vignettes, take the audience's (club members') perspective (think about their expectations, etc.), and write whatever came to their minds. In contrast, the participants in the control group were asked to read short vignettes but were not explicitly requested to take the audience's perspective. They were just asked to write whatever came to their minds. Then, both groups were given 40 minutes to share their written ideas about allocating a limited amount of money among the student clubs (see Materials for more details).

Importantly, besides the experimental manipulation, EF (high versus low) and trait PT (high versus low) functioned as independent variables. The main questions were whether the experimental task would increase these novice EFL writers' performance (overall writing quality and the level of persuasiveness of the passages) and whether EF or trait PT would moderate this effect, and also whether trait PT would compensate for the lack of EF resources. The primary findings confirmed the efficiency of the experimental task and its interaction with these psychological variables under certain circumstances.

I will summarize and discuss these observations in the following section after shortly concluding the manipulation check analyses. Then, I will speculate how mental

writing processes might have been involved/integrated, particularly when the participants in the experimental group were exposed to an audience-oriented task. Next, I will further discuss the study's observations and suggest future directions. Lastly, pedagogical implications, the study's possible contribution, and limitations will be shared.

7.1 Conclusion

In this section, I will first consider the success of the experimental manipulation. Following that, I will focus on the conclusions related to the overall writing quality by considering the main questions and the first, the third and the fifth hypotheses of the study. Then, I will focus on the conclusions related to the level of persuasiveness dimension by considering the main questions and the second, the fourth and the sixth hypotheses of the study. After that, I will speculate the possible integration of writing processes on the basis of the observations of this study and a cognitive writing processes model. I will finish with an overall conclusion.

7.1.1 Manipulation check

As mentioned before, this study employed four different manipulation checks, considering before and while writing the main passages as well as participant reports and more objective evaluator assessments (e.g., coding). All of them indicated that the manipulation was successful at orienting the participants to the audience in the experimental condition.

First, the participants in the experimental group reported that they tried to take the audience (club members) perspective more when they read the messages presented in vignettes before writing. Second, they reported that they considered the audience's perspective more than the control group during the main writing. For example, compared to 25 participants in the control group, only 4 participants in the experimental group reported no PT while writing. And importantly, among those who took the audience perspective, compared to 27 participants in the control group, 49 participants in the experimental group reported that taking the audience perspective influenced their writing processes (e.g., organization or problem-solving).

Third, the evaluation of the three-minute reaction notes taken before writing, during task processing (PT in the experimental and a kind of brainstorming in the control condition), indicated that the experimental group attempted to take the audience's perspective more frequently. Lastly, and maybe most importantly, when the EMs which may be "the clearest signals that a writer is considering the presence of an active audience" (Hyland & Jiang, 2022, p.6), and SM markers which may reveal the most potent self/author representation (Hyland, 2001; Ivanic, 1998) and which may be the indicators of PT (Galinsky et al., 2005) were considered, the analyses revealed an EM versus SM marker dissociation in the main writings. And this was much more visible in the experimental than in the control condition. More specifically, the mean values of almost all EM subcomponents were higher, whereas SM subcomponents were lower in the experimental compared to the control group. Additionally, the analyses revealed that inclusive EMs were the most frequently employed EMs in the experimental condition. Thinking that PT activity was suggested to create a merging of

self and other (Davis et al., 1996), one possible interpretation is that the participants in the experimental group tended to put themselves into the audience's shoes more.

Overall, these four measurements, which aimed to check the success of manipulation, confirmed that the task could orient the participants to engage with their audience during task processing and while writing. Particularly, the higher frequency of the employment of EMs may indicate that the audience was relatively maintained more in the mind of the participants in the experimental group while they were writing the main experimental passages.

7.1.2 The influence of the task, EF and trait PT on the overall writing quality

This study's first and third hypotheses predicted that exposing novice upper-intermediate EFL writers with high EF or trait PT to an audience-oriented task would increase their overall writing quality performance. Additionally, the fifth hypothesis predicted that if the high trait PT contributed to writing performance, it could compensate for the lack of sufficient EF resources. The results provided clear support for the first and the third but only partial support for the fifth hypotheses. I will now summarize and discuss these findings.

The analyses indicated that EF (low versus high) can interact with the task (audience-oriented experimental versus non-audience-oriented control). The results revealed that the participants with relatively higher EF performed better than the ones with relatively lower EF, but more importantly, the experimental (audience-oriented) task distinctively increased the performance of participants with relatively higher, rather than lower EF capability. The difference (~5 points) between low and high EF groups

was doubled (~10 points) when this high EF group was exposed to the experimental task. Furthermore, the analysis revealed that the effect of the experimental task for this high EF group remained significant even after controlling/fixing the participants' language background and trait PT. Hence, the findings confirmed the central role of EF resources in writing and indicated that encouraging to actively engage via an audience-oriented task can differentially boost the performance of novice upper-intermediate EFL writers with higher EF, who have sufficient resources for responding to the task's requirements.

Regarding trait PT, the analyses revealed its interaction with task conditions, as well. The source of this interaction was the better performance of the participants with relatively higher trait PT in the experimental condition. Furthermore, similar to EF, the differential effect of the experimental task for the high trait PT group remained significant even after controlling/fixing the language background and EF. Hence, although it may not influence the writing performance by itself, the high trait PT may be activated by an audience-oriented task and boost the novice EFL writers' performance.

As mentioned before, one potential of having high trait PT may be compensating for the lack of sufficient EF resources to deal with the extra demands which may emerge from active audience processing. Focusing on the audience-oriented experimental condition and comparing the performance of the low EF – high PT and low EF – low PT groups, which differ only in their trait PT abilities, could be critical to answering whether this was the case. If the low EF – high PT group performed similarly to the high EF – high PT and the high EF – low PT groups but performed better than the low EF –

low PT group in the experimental condition, it could indicate that trait PT may compensate for the audience processing demands on EF resources.

The analyses did not reveal statistically crystal-clear conclusive results in this regard. More precisely, the descriptive mean values indicated that the mean overall writing quality performance of the low EF – high PT group was in between. It was higher than the low EF – low PT group but also lower than the high EF groups (see Table 29). Regarding the inferential analyses, although the difference between low EF – low PT and low EF – high PT groups were not significantly different, it was revealed that the low EF – high PT group's performance was not, but the low EF – low PT group's performance was significantly lower than the high EF groups. Based on these data, because the performance of low EF – high PT was not significantly higher than low EF – low PT group, I accept that this conclusion should be approached with caution, but I tend to conclude that the high PT may be compensating for the limited cognitive EF resources to some extent and approach the low EF – high PT group's performance to high EF groups.

Apart from these main investigations, seeing that the experimental task may have a detrimental effect on the low EF – low PT group, I focused on them and conducted further additional analyses to better understand the role of the tasks for this group. The transformed z-score comparisons across the diagnostic and the experimental main overall writing quality values revealed that although these participants performed similarly in the diagnostic writing, exposing them to the experimental task significantly decreased while exposing them to the control task significantly increased their passage's overall writing quality. Thus, it seems that exposing those low EF – low PT participants

to an explicit audience-oriented task may not be so beneficial⁶³. When it comes to the reason behind this pattern, thinking that the most crucial difference between the experimental and control conditions was the existence of short state PT versus brainstorming activities of the tasks, simply asking them to brainstorm and write whatever came to their minds may be the reason for approaching their performance to the sample's mean ($z = 0$) (see below for further discussion)⁶⁴.

Beyond these causal effects, additional correlational analyses revealed a positive association between the diagnostic overall writing quality and EF capability but not between the diagnostic overall writing quality and trait PT. However, further investigation revealed that the latter association emerged when the passages written in the main experimental (but not in the main control) condition were considered. Hence, the correlation between the overall writing quality and trait PT may be dependent on the exposure to the audience-oriented task. As also concluded from the regression analyses results (see above), trait PT may be activated and associated with the overall writing quality when the attention of those upper-intermediate novice EFL writers was directed to the audience.

The findings of this study follow the literature which emphasized the involvement of EF (Vasylets & Marin, 2021) in different writing processes such as the

⁶³ This was not the case for the low EF – high PT group. That is, the audience-oriented task did not worsen their performance, which may be another tentative, although insufficient, support for high-trait PT's compensation for the lack of EF resources.

⁶⁴ Still, further studies are surely needed, for example, to isolate brainstorming activity and investigate its effect before concluding.

formulation (planning and translating/formulating⁶⁵), execution (programming and executing/transcribing) and evaluation (reading and editing) of ideas (Kellogg, 1996). They also follow the literature that emphasizes the importance of PT (Cho et al., 2021), which is associated with audience awareness (Kim & Park, 2019). Overall, the findings are in accordance with, for example, the DIEW model, which suggested that domain-general cognitions (EF) create a basis for both the low-level writing processes, such as the employment of foundational language, and high-level regulatory processes, such as PT which can determine the level of writing performance (Kim & Graham, 2021). In the following parts, I will further associate the literature with this study's data and make conclusions.

The importance of considering the audience who is not immediately available (e.g., Bereiter & Scardamalia, 1987) and the importance of designing audience-oriented writing tasks (e.g., Cohen & Riel, 1989), particularly for novice writers, were recognized by many researchers in the area (Block & Strachan, 2019; Cho & Choi, 2018). This was because “generally speaking, writing is an activity designed to create a text for some audience” (Hayes, 2012, p. 375), but novice writers who may still be struggling with fundamental language processes (McCutchen, 2000) and “lack experience in understanding reader's needs, and thus remain very writer oriented as they review their text” (Becker, 2006, p. 47) may ignore their audience (Hayes, 1980b). Fortunately, some studies indicated that although it may recruit the limited cognitive resources (Roxßnagel, 2000), directing the attention to the audience may help the

⁶⁵ Importantly, similar to the researchers in the area (e.g., Zimmermann, 2000), in this chapter, I will use the root *formulate* and its several word forms, rather than the root *translate* and its variations, to avoid confusion with the translation related to L1 to L2 processes.

participants to take other's (audience) perspective better (Apperly et al., 2006) and creating an appropriate reader-oriented task can improve their writings (Traxler & Gernsbacher, 1993). The results of this study confirmed and extended these observations.

The designed audience-oriented experimental task in this study successfully directed the attention of novice EFL writers to their audience and increased the overall writing quality. However, the task's success depended on the availability of particularly EF resources or trait PT tendencies. Hence, it is crucial to understand the nature of these resources and their involvement in writing processes to be able to answer how they moderated the audience-oriented task.

Regarding EF, it may be involved in language and writing processes in many different ways (Kellogg, 2001; Kim & Park, 2019), such as in inhibiting/suppressing irrelevant ideas or less sophisticated lexical items while formulating (Olive, 2011), or suppressing L1 while writing in L2 (Li, 2023). And as the Capacity Theory of Writing (McCutchen, 1996) suggested, the availability of limited EF resources may determine/predict the successful employment of writing processes such as the retrieval of the lexical items from LTM, text generation (fluently turning ideas into text) (McCutchen et al., 1994) which may be reflected in the produced syntactic complexity (e.g., longer utterances) (Kemper, Kynette, Rash, O'Brien, & Sprott, 1989), and the overall structural text complexity (e.g., a hierarchy of the embedded events) (Hoskyn & Swanson, 2003). Thus, it is evident that writing, which is a productive skill that includes language processes such as idea generation, formulation (in oral language), and transcription, would require limited EF resources (Kim & Park, 2019).

It was also reported that EF's involvement might be more observable for novice writers (McCutchen, 2011), particularly if the language generation processes are not fluent enough (McCutchen, 2000; Schmidt, 1992). In other words, language proficiency may be one of the most important factors which interact with and consume EF resources (Schmidt, 1992), and this may have far-reaching consequences such as ignoring the audience (Hayes, 1980b) or adapting, for example, knowledge telling strategies by identifying some topics (e.g., theatre, outdoor sports, and music in this study's context), and using the topic identifiers (e.g., theatre) as memory cues to automatically activate, retrieve and employ the related concepts without monitoring or planning (Scardamalia & Bereiter, 1987). Using these kinds of strategies may work to some extent and reduce the load on the limited EF resources (McCutchen, 2011). However, the focus of the EFL learners on non-automatized (controlled) lower-level processes, such as the language form (Schmidt, 1992), rather than the higher-level processes such as text cohesion (Knospe, 2017) may decrease the quality of writing (Kobayashi & Rinnert, 2018).

As reported before, besides fundamental language processes, the higher order processes such as PT which is a high-level inferencing activity (e.g., making inference about the mental states of others), is essential for writing development and is involved particularly in higher quality writing (Kim & Park, 2019). One aspect of PT taking is audience awareness (Cho et al., 2022), and it is suggested that all kinds of texts need to orient toward their intended audience properly, and this process needs to take potential readers' perspective, including their expectations, needs, and background knowledge (Grabowski, Mathiebe, Hachmeister, & Becker-Mrotzek, 2018). However, active audience processing, such as taking the other's perspective, may also require limited EF

cognitive resources (Davis et al., 1996; Kim & Park, 2019; Roxßnagel, 2000). Hence, the participants who performed well in the EF task may perform better in language production by considering the audience perspective (Ryskin et al., 2015), which may be, for example, because they may be adapting their utterances by considering the audience perspective (Wardlow, 2013) depending on the aims of the conversation (Yoon, Koh, & Brown-Schmidt, 2012). Thinking that the EF resources are required by the controlled (more than the automatized) processes and having a natural tendency to consider other people's situation may help the writer to process the audience (expectations, etc.) more automatically, having higher trait PT may ease some portion of these resources. This may be one of the reasons behind the success of the writers who take the audience perspective better (Cho et al., 2021), for example, by considering their linguistic and cognitive competence (Roxßnagel, 2000) as well as their needs or attitudes towards the topic (Weigle, 2005), and make some finer decisions (e.g., in word choice) (Phrases, 2002) and adapt the content to create better passages (Midgette et al., 2008).

To repeat once again, the results of this study summarized above confirmed this literature. They revealed that an audience-oriented task might be beneficial for the higher EF group or higher trait PT group, and although it was not quite clear, the higher trait PT may compensate for the lack of sufficient EF resources.

When it comes to how EF and trait PT might be involved within this study's context, based on the findings and the briefly reviewed background, I conclude that, apart from the others, both non-fluent language processes and audience-oriented processes which were successfully triggered by the experimental task demanded the limited EF resources of those novice upper-intermediate EFL writers. Activating

audience-oriented processes seemed more beneficial for those with relatively high EF resources probably because they could manage these demands more efficiently and benefit from maintaining the audience in their minds. The increased overall quality scores indicate that they could maintain audience in their focus and integrate it into the other processes more successfully.

Again, the activation of audience-oriented processes was beneficial for those who have relatively high trait PT, probably because they may process the audience with less cognitive effort and more efficiently and, for this reason, audience processing while writing may be more manageable for them. Their overall writing quality scores also show that they could integrate the audience into the other writing processes and, at least to some extent, successfully orchestrate them. Besides, there is a possibility that trait PT can compensate for the lack of sufficient EF resources to deal with the audience-oriented task demands by probably decreasing the load of audience processing.

However, the audience-oriented task may not be so beneficial, particularly for those who do not have advantageous EF capability and trait PT tendencies. As shared above, the additional analysis revealed that the task which encouraged keeping the audience in mind while writing might overload the limited resources and have some detrimental effects on the overall writing quality of those participants who have neither relatively high EF capability nor trait PT tendency (low EF – low PT). Regarding how this effect might have occurred, it may be necessary, at this point, to note that the investigation of self-reports and the frequency of EM and SM markers revealed that this group's audience processing pattern was similar to the other experimental groups (e.g., high EF – high PT). In other words, like the other experimental groups, this group was

actively engaged with the audience during writing. This may be related to this group's low performance in the experimental condition because the active engagement might interrupt their default allocation of EF resources, for example, for the formulation processes. As mentioned above, both micro-level (e.g., fundamental language) and macro-level (e.g., audience) processing may compete to recruit EF resources. And although the writing models (e.g., DIEW) focus on and emphasize the probability of the EF resources taken away by the relatively lower-level fundamental language processes, anything which loads the system may interrupt with, for example, formulation processes (e.g., Kellogg, 2001). Thus, the influence of the load may work vice versa: that is, the higher-level processes may also create a constraint for the lower-level processes (Flanigan, 2021), and this constraint might have influenced the overall writing quality of this group.

Actually, although it was not the main concern of this study, the additional analyses revealed that exposure to a relatively more straightforward task, which includes a simple brainstorming rather than state PT activity, may increase and approach this group's performance to the overall average of the study sample. Again, the explanation for this observation may be related to the active employment of the writing processes and the allocation of EF resources. As Kellogg (2001) indicated, EF resources may be simultaneously recruited by planning, formulating, and reviewing activities. But the task requirements such as focusing on the language form versus content (Knospe, 2017) may create a competition between them (Skehan, 1998, 2009). And particularly if the EF capacity is not high enough, the system may have difficulty in integrating the processes and formulating the ideas (McCutchen, 1996).

However, as Kobayashi and Rinnert (2018) reported, generating ideas (e.g., retrieving information from LTM via a freewriting/brainstorming activity) before starting to write may contribute planning process (Kobayashi & Rinnert, 2018), which may free some of the limited EF resources for the other competing processes, such as formulating or reviewing (Kellogg, 2001). This effect may be much more pronounced for those with EF resources just sufficient to maintain even the most fundamental writing processes simultaneously. This may explain why the control task (including a kind of simple brainstorming) was more beneficial and pronounced for this low EF - low PT group. Being exposed to such a task, which does not explicitly increase the existence of the audience in their consciousness, might have helped them, for example, to create a writing template beforehand and ease some of their limited EF resources for the other processes, such as generating ideas while writing.

7.1.3 The influence of the task, EF and trait PT on the level of persuasiveness

This study's second and fourth hypotheses predicted that exposing novice upper-intermediate EFL writers with high EF or trait PT to an audience-oriented task would increase their level of persuasiveness performance. Additionally, the sixth hypothesis predicted that if the high trait PT contributed to writing performance, it could compensate for the lack of sufficient EF resources. The analyses conducted for the overall writing quality were also conducted for the level of persuasiveness. The results provided only tentative support for the second hypothesis. The participants with relatively higher EF capability seemed to perform better in the audience-oriented

experimental task, but the difference was only marginally significant. The other hypotheses were rejected. I will now summarize and discuss these observations.

The existence of a marginally significant interaction between the tasks and EF and the lack of any interactions between the task and trait PT or their main effects may be due to the challenging nature of persuasive passage writing (Varghese & Abraham, 1998), which might have created a kind of floor effect for those novice EFL writers. Writing these passages is “an intellectually challenging problem” that requires processes, such as anticipating alternative perspectives and critically evaluating arguments, which may be, for example, “constrained by the limitations on processing capacities and the tendency to privilege one’s perspective (my-side bias)” (Ferretti & Fan, 2016, p. 311-12). Probably because of this, formulating written arguments to convince the audience is quite difficult, even for native speakers of English (Ferretti & Fan, 2016; Gilbert, 2004). As would be expected, the difficulty may be more severe for EFL learners (Qin & Karabacak, 2010; Qin & Liu, 2021; Silva, 1993) and particularly for those whose proficiency level is relatively low⁶⁶ (Cheng & Chen, 2009). The data of this study provides a support for this suggestion because although the average overall writing quality performance of the sample was not very bad (68.6 out of 100), the average level of persuasiveness performance was quite low (just 3.36 out of 7). Hence, there was a kind of floor effect when the level of persuasiveness dimension was considered.

One reason for this may be the relatively low L2 proficiency processes of these EFL writers, which may consume their limited cognitive (EF) resources, and which in

⁶⁶ For example, vocabulary knowledge can predict the development of reasoning (Richland & Burchinal, 2013).

turn may create a challenge in managing the other cognitive demands (Kobayashi & Rinnert, 2018) for showing argumentative competency (Osborne et al., 2016). More specifically, the cognitive systems of those novice EFL writers may not deal with the other simultaneously demanding processes because of their not-sufficiently proceduralized L2 language processes (Schmidt, 1992). They may be too busy to work on expressing their ideas on a micro-level, and because of this, although they may improve their passages' overall writing quality, they may not sufficiently focus on macro-level concerns (e.g., reasoning), which may contribute to highly demanding (e.g., compared to the narratives) persuasiveness (Kellogg, 2001; Kobayashi & Rinnert, 2018).

According to Kellogg (2008), these types of persuasive passages do not only include “what to say” and “how to say” but also reasoning processes that may require the writer’s limited cognitive resources (EF). Moreover, these reasoning processes themselves may require coordinating many cognitive processes, such as evaluating the underlying assumptions (Eemeren et al., 2002). Actually, EF may be so important for reasoning (e.g., for creating persuasive passages) that early-age EF capability may predict its development (Richland & Burchinal, 2013) and, as some researchers suggested, reasoning and problem-solving (or fluid intelligence) may be considered critical components of higher-level EF processes (Diamond, 2013). The extensive analyses of these processes are beyond the scope of this study; however, it seems that following sound reasoning and creating convincing passages may require quite high EF resources.

This may be one of the underlying reasons behind the differential observations regarding the overall writing quality (see above) versus the level of persuasiveness dimensions of the written passages. The audience-oriented task may help those relatively low proficiency level novice EFL writers to deal with and improve the overall writing quality (see Appendix H for the rubric), but it cannot (at least by itself) help those writers work on and improve the passages' level of persuasiveness by introducing convincing reasons (see Appendix I for the rubric).

Still, it is important to notice that this study revealed some clues that signal EF capacity's importance in creating convincing passages. Although it was at a marginal level, as the interaction analysis revealed, when the attention of those who have higher EF capacity was directed to the audience via the experimental task, they could deal with the overload better and could increase their passage's level of persuasiveness⁶⁷. Thus, these novice EFL writers obviously have difficulty creating convincing passages/arguments. However, although the audience-oriented task may create an extra load by keeping the audience in focus, it may be beneficial for novice EFL writers who have some EF resources to allocate for this process. In other words, the experimental task may potentially increase the persuasiveness of the passages written by those with relatively higher EF resources. This seems to confirm that EF is a critical component in this equation and deserves further consideration.

⁶⁷ Additionally, although the difference was not statistically significant and the mean difference was small, the pattern was reversed for the low EF group ($M = 3.32$ versus $M = 3.06$). Their mean performance seemed to be better in control than in the experimental condition. Although I tend to interpret this as the absence of the load created in the experimental condition (audience processing while writing), the statistics are not strong enough to make conclusions, and this was not the main concern of this study.

Lastly, when it comes to the involvement of trait PT, the data revealed that having high trait PT does not contribute to the level of persuasiveness dimension. Beyond but also related to the floor effect created by the challenges of persuasive writing, one additional tentative explanation for the lack of this and the other effects with respect to the level of persuasiveness may be related to the timing of the manipulation. In this study, I gave the audience-oriented task before the participants started to write the main experimental passages. This was appropriate for the aim of the study, investigating the involvement of high-level audience processes as they compete with low-level (particularly non-automatized fundamental language) processes for the limited EF resources. Furthermore, many studies indicated that increasing audience awareness before writing may improve writing performance (e.g., Block & Strachan, 2019; Cho & Choi, 2018; Cohen & Riel, 1989).

However, some other studies, for example, Roen and Willey (1988), revealed that although introducing audience awareness manipulation both before writing (e.g., asking about the prior knowledge of the audience) and before revising (after writing the first draft) contributed⁶⁸, introducing the audience manipulation before the revision stage was more beneficial ($M = 6.75$) than introducing it before writing the first draft ($M = 6.00$). Some other studies also confirmed this boosting effect of introducing the audience manipulation before the revision stage. For example, Traxler and Gernsbacher (1993) revealed a positive reader PT effect on written passages when the reader's perspective was taken before the revision (after writing the first drafts) of descriptive essays, and

⁶⁸ The results of their study revealed that overall audience manipulation increased the quality of the passages: The participants performed worse in the no audience ($M = \sim 5.35$) than in the audience manipulation conditions (see above for the other mean values).

Midgette et al. (2008) indicated an increased argumentative writing performance when the participants were directly asked to think about the reasons of possible opposing readers, and when this was done before the revision stage. Again, Wang (2016) indicated that adding an interactive audience task before the revision rather than before writing the first draft (during planning), although slightly, increased the participants' persuasiveness performance.

This could be expected for two reasons. First, many writing processes (e.g., idea generation and planning) are highly active while writing an initial draft and consume EF resources, but these processes are probably less active and the limited EF resources could be eased while writing for revision purposes (Kellogg, 2001). Second, the main focus of many studies which revealed this effect was on persuasiveness (e.g., Midgette et al., 2008; Wang, 2016) which is directly associated with audience awareness or PT processes such as considering the opposite arguments (Reznitskaya et al., 2001) and may require high EF resources. This extra demand may not be met by the limited EF resources, particularly when the other writing processes are highly active during the first draft. This may create a floor effect and make the effect of the audience-oriented task manipulation invisible, which may explain why I found no main effects of the task, trait PT, and EF, and only a marginally significant interaction between the audience-oriented task and EF. The task's influence was restricted to high EF participants. Based on these, I suggest that introducing an audience-oriented task after the novice EFL writers compose the first drafts but before the revision stage would pronounce its effect more because they may allocate more resources and activate more audience-oriented and

reasoning processes that may interact with, for example, trait PT to create a convincing passage.

7.1.4 Speculation of the possible integration of mental writing processes

Although this study did not aim to evaluate any of them, and it is not possible to completely fit the data, it may be explanatory to focus on one of the cognitive writing models briefly and try to get insight into the possible mental processes of the experimental group by considering their EF capability and trait PT tendencies. This may be done partially on the basis of the analyzed data and partially by speculating on the basis of the literature and the observations. To remind it again, similar to the researchers in the area (e.g., Zimmermann, 2000), I will continue using the word *formulate* rather than *translate* to avoid confusion with the translation concept related to L1 to L2 processes.

For this purpose, I will focus on Hayes's revised cognitive writing processes model (Hayes & Berninger, 2014) and try to understand the mental processes of the participants in the experimental group. I may also consider and compare the control group if needed. Because data is clearer on overall writing quality performance dimension, the focus will be on it.

As mentioned in the 2nd Chapter, this model includes three (resource, process, and control) levels. Most critically, the process level of the model (Hayes & Berninger, 2014), which includes the task environment (collaborators & critics, transcribing technology, the text written so far, and task materials) and writing processes planes (proposer, formulator, evaluator, and transcriber) may be controlled by the interaction

between the components within itself and between the other, resource and control, levels. Although I will simplify my speculation, the relationships among the levels and individual components may be non-linear and reciprocal. The processes or the components from the model will be italicized.

To remind once again, the current study aimed to create an experimental task that may direct the attention resources to the audience and help them manage the writing processes more efficiently (e.g., by considering the audience while formulating meaningful sentences). The results indicated that it worked for those with sufficient EF resources to deal with active audience processing or those with high trait PT who tend to consider the other people's needs, expectations, etc. The latter may interact with the triggered audience processing and ease the cognitive load. Thus, it may compensate for the lack of EF resources.

When I consider how the experimental task influenced cognitive writing processes, one possible scenario is as the following. The *task initiator* (the researcher) successfully introduced the *task materials*, which employed audience-oriented (e.g., state PT) components and encouraged active audience processing. Following this introduction, the audience (and other relevant) information from *LTM* was retrieved. This interaction between the *resource* and *control* planes triggered the involvement of the *collaborators & critics* component and brought the audience (e.g., club members) psychologically closer⁶⁹ (or even immediately available). The level of audience

⁶⁹ The research indicated that taking the other's perspective (particularly familiar – concrete and real - ones) may increase the psychological closeness between the perspective taker and the other (Smith, Soderberg, Netchaeva, & Okhuysen, 2023) and make these other's perspectives more cognitively accessible which may result, for example, from an overlap in self and other mental representations (Ku et al., 2015).

representation details may be dependent on trait PT ability. If the writer had a higher trait PT tendency, a more detailed picture of the audience may be visualized, and this can be done relatively more easily. Importantly, the audience representation was probably taken into focus and kept/updated during writing by the limited *EF* resources⁷⁰, which were also probably used to shift the perspectives among the audiences (e.g., from the rector to the music club members). Once a different perspective is taken, the others might have been suppressed. This can be the other stage where trait PT was also involved, for example, by changing perspectives more efficiently and easing the load on *EF* resources. Additionally, the *planner* was probably employed to organize the task fulfillment while, for example, keeping the audience in mind, and this process recruited additional *EF* resources.

Thinking that the curriculum of the preparatory school did not specifically teach genres and the participants were novice EFL writers, I tend to think that the *genre knowledge* embedded into the *writing schemas* (writers' beliefs about the target properties of a text) did not, but the *strategic knowledge* (how to produce a text) significantly contributed to controlling the writing processes. Other than that, in this study's context, the other individual resources, *reading ability*, and the *LTM* with respect to topic and audience familiarity backgrounds were controlled. However, depending on the successful integration of the audience processes, they might still have interacted with the tasks, and, for example, the participants in the experimental group

⁷⁰ See the Manipulation Check Analyses section for the evidence (e.g., decreased employment of SM markers) that the task might have successfully modulated the writing processes by directing and keeping the experimental group's attention on the audience.

might have read *the text written so far* more frequently and retrieved more audience-relevant background information from their *LTM*s⁷¹.

When it comes to the reflection of the experimental condition on more specific writing processes, as I mentioned above, the *task initiator* and *the task materials* probably directed the attention to the audience, recruited the limited *EF* and the relevant *LTM* resources (e.g., for adopting strategies and taking audience perspective), activated the *collaborators & critics*, employed the retrieved *writing schemas* (strategy use), and used the *planner*, etc. To speculate further, the activated *proposer* got input from the *collaborators & critics*, *the task environment*, the *planner*, the *text written until that moment*, *writing schemas* (strategy use), the *LTM* and audience processes, and made some nonverbal suggestions for the text. It is likely that the interaction of the *proposer* with the other components influenced its processes. For instance, the *evaluator* was involved more or less depending on the strategy used and the availability of *EF* resources and interacted with the *proposer* by checking its suggestions more or less frequently. The *formulator*, which received the proposed ideas and converted them into verbal representations (which probably were simultaneously checked and finalized by the *evaluator*), conveyed them for the *transcription*. Importantly, the participants exposed to the audience-oriented task probably kept the audience more in their focus and evaluated the proposed ideas more efficiently before *formulating* and *transcribing*. And this might be the case, particularly

⁷¹ Although the familiarity with the club activities was controlled (e.g., choosing a familiar topic after piloting alternative options and presenting vignettes) and the analyses revealed that they did not influence the writing performance in this study, as Fayol (2016) reported, under normal circumstances, because the topic knowledge may be easily accessed, being familiar or having domain-general knowledge may ease the cognitive load and improve the text's quality via, for example, formulating processes.

for participants with enough *EF* resources. Other than that, thinking that these novice EFL writers had relatively low language proficiency levels, the *formulation* process might have recruited more *EF* resources and created a bottleneck effect (see below for more details). The participants who were in the experimental group and did not have enough *EF* resources might have had to sacrifice, for example, from audience or language processes. When it comes to those with higher *EF* or higher trait *PT*, the bottleneck effect's constraints might have been reduced, and the *formulator* relatively more easily realized the formulated ideas as written texts. Lastly, because they need to keep the previously written/transcribed passage actively in their minds as they progress, the load on the *EF* resources might have been increased as they continue writing, and their performance worsened as they approached close to the end.

To emphasize once again, although I simplified these writing processes to present them more clearly, they are primarily non-linear. For example, the *evaluator* might be active before, after, or while *formulating* or *transcribing* and may modulate the *planner* and *proposer*. It may also be essential to note that, as these processes more or less interact and work simultaneously, it is likely that, particularly those control group participants with lower *EF* adopted a kind of stream-of-consciousness writing strategy to deal with the domination of limited *EF* resources by non-fluent language processes⁷². However, those with lower *EF* capability in the experimental condition might not have ignored the audience maintained in their minds.

⁷² For example, as would be in accordance with knowledge telling strategy (Bereiter & Scardamalia, 1987), the topic assigned by the task could function as a cue and have helped those novice writers who have low *EF* and low trait *PT* to keep track of the content and to retrieve readily available knowledge from their LTM to deal with, for example, the *EF* based constraints (McCutchen, 1996). Still, this may not mean that the *proposer* needed to keep the general topic in the focus of the writer (Hayes & Berninger, 2014).

As mentioned above, this extra load probably interrupted, for example, their low-level (e.g., *formulation*) processes and worsen their performance. However, it is likely that those with higher *EF* resources benefitted from this active audience processing by orchestrating their resources better. Although the exposure to the audience-oriented task, which brought the audience (club members and the rector) information from their *LTM* to their attention, might have resulted in the extra demands from *EF* resources, this group seemed to balance the allocation of focus on low-level and high-level writing processes and integrated the audience into their writing processes more successfully. Additionally, it is possible that, again, as mentioned above, participants with more experience or natural tendencies to take other people's perspectives (trait *PT*) might have experienced the active audience processing load less, which might have helped them manage their resources somewhat similar to the high *EF* group.

Hayes' revised cognitive writing processes model (Hayes & Berninger, 2014), which I used above to speculate the possible orchestration of the processes focusing on the participants in the experimental group, was successfully extended into the EFL context (e.g., Kim, Tian, & Crossley, 2021). Additionally, some researchers suggested that the difference between L1 and L2 processes is mainly quantitative (Silva, 1993) rather than qualitative (Wang & Wen, 2002).

However, *proposing* and *formulating* processes that may be influenced by L1-L2 interaction more may deserve a closer look. This is because these processes, which may differentiate the more versus less skilled writers (McCutchen et al., 1994), may be essential processes which consume *EF* resources (Olive, Kellogg, & Piolat, 2008), particularly for EFL learners (Li, 2023). In the following parts, based on the restricted

data collected in this study and some findings in the literature, I will argue that low proficiency novice EFL writers may benefit from L1-L2 translation during *proposing* (idea generation); however, particularly formulation of these ideas, *formulating* process, may create a kind of bottleneck.

In this study, although it was not the main aim and I did not systematically investigate it, I collected some data which may help me get a little insight into *proposing* and *formulating* processes of those upper-intermediate level novice EF writers. First of all, almost all the participants reported that they depended on their L1 for thinking (idea generation) (see Language Background Part for details), which may mean that the concepts were proposed in L1 Turkish and then formulated in L2 English before or while transcribing⁷³. Additionally, when I got in touch with some participants and asked them whether thinking in Turkish and then writing in English was beneficial, most of the 19 responders reported that it was beneficial for several reasons, such as feeling psychologically more comfortable, thinking more deeply, analyzing better, creating better sentences in L1, and writing more fluently⁷⁴. Two of those students who reported benefits also reported that they experienced some difficulties, for example, in translating

⁷³ This is in accordance with the observations in the literature. Although some studies suggest that the influence is minor (e.g., Zimmermann, 2000), many others, as reported above, suggest that there is a significant influence of L1 on L2 with respect to formulating processes (e.g., the ideas generated in L1 and then translated in L2) which may be particularly the case for the EFL writers whose language proficiency is lower (e.g., Wang & Wen, 2002).

⁷⁴ This observation is also in accordance with some empirical studies. For example, Kobayashi and Rinnert (1992) investigated the translation (first writing in Japanese and then translating into English) and direct writing performance of lower versus higher proficiency EFL writers and found an interaction effect. The L1-L2 translation resulted in higher quality writing and most importantly the lower proficiency participants, rather than the higher proficiency ones, benefitted from the translation condition. There may be many reasons for benefitting from this kind of a process. For example, because of the lack of fluent L2 retrieval system, they may not be able to efficiently use their WM to access their L2 content in LTM (Weigle, 2005) and as a result they may be benefitting their L1 resources, for example, for planning (e.g., generating ideas) or (even) formulating (e.g., word searching) (Kobayashi & Rinnert, 2018).

complex sentences, and they lost time. Other than these, five students who retrieved ideas in L1 Turkish and translated them into L2 English reported difficulties, for example, in keeping the complexity of the original ideas or forgetting them. To note, those participants' EF capability⁷⁵ was close to the overall average. Beyond these, this restricted data did not reveal an important difference between the experimental and control groups.

On the basis of this data, it seems that the L1-L2 translation strategy may be influencing idea generation and formulation processes somewhat differentially. Probably because trying to retrieve/generate and work on all the ideas in their L2 English could be more difficult (even impossible) for most of those students who did not have enough L2 language resources, although some of them reported experiencing difficulty, which resulted from translating ideas (e.g., forgetting), they continued using L1 Turkish to generate them. More specifically, they seemed to be dependent on L1 because they did not have enough L2 resources, L1-L2 translation made them feel safer, and it helped them at least to generate and work on the ideas more easily. Thus, based on their reports (e.g., thinking more deeply), I tend to think that these participants may benefit from L1 during idea generation and proposing⁷⁶.

⁷⁵ My tentative interpretation on the basis of this limited data is that although using L1 while writing in L2 may consume EF resources, the use of this strategy is not necessarily dependent on the EF capability for this upper-intermediate level novice EFL writers. That is, it is not only those participants with low or high EF capabilities who adopted this strategy.

⁷⁶ Still, it may be important to emphasize that I did not systematically investigate the language used to generate ideas, for example, via think-aloud protocol while the participants were writing the passages in this study. Thus, this suggestion must be approached tentatively. Future studies may investigate whether such a task influences the employment of L1 or L2 while writing.

Still, it may be important to note that although generating ideas in L1 may be functional for this level of students, it does not mean that it does not lead to extra loads. For example, as Nawal (2018) indicated, generating ideas in L1 may split attention, create a cognitive load, and consume more of the limited resources (e.g., because of the need to mentally shift between the languages, or simultaneous attention on both of them, etc.) which may prevent the EFL writer, for example, from conveying meaning more smoothly by employing appropriate sentence structures and L2 vocabulary. Hence, although, for example, L1-L2 translation may have some benefits, depending on the language proficiency (e.g., less linguistic experience may mean more frequent revisions) and the availability of WM (EF) resources (Chenoweth & Hayes, 2001), particularly the formulation of those ideas may create a bottleneck that limits the fluency of the system which conveys the ideas retrieved from the proposer or LTM (Hayes & Berninger, 2014).

This suggestion seemed to have a support from the student reports which conveyed some eye-catching problems such as the difficulty of formulating the generated ideas in L2 while keeping its original complexity (or even forgetting them). Furthermore, as some researchers reported (e.g., Li, 2023), for example, insufficient linguistic resources in L2 writing may pose greater cognitive demands and because the more effortful and conscious formulating in L2 may consume more EF resources, the system may be overloaded while writing in L2. It may be so demanding that according to Wang and Wen's (2016) study, which collected writing processes data via think-aloud protocol data, two-thirds of EFL writer's⁷⁷ attention may be allocated to sentence

⁷⁷ The data were collected from intermediate and advanced level students, but it was not specified for which level this conclusion for. However, because they also reported that the lower proficiency EFL

construction, less than one-third to idea generation and only a little for the organization. Thus, compared to the others, it seems that *formulating* (or text-generating) processes may be quite difficult and resource consuming. On the basis of these restricted observations, my tentative suggestion is that L1-L2 translation strategy may have some advantages particularly during the idea generation. However, formulating sentences in L2, particularly because of not sufficiently proceduralized language processes, may create a kind of bottleneck⁷⁸ (e.g., while retrieving words and creating sentences in L2) which results in the loss of original ideas.

Turning back to my speculation about the possible writing processes of the experimental group on the basis of Hayes' model (Hayes & Berninger, 2014), I suggest that writing in L2 might have created a difference, particularly with respect to the extra load that emerged by the employment of L1-L2 *formulation* processes which may need to activate and deactivate its different levels, for example, during vocabulary or sentence structure choice. This might have interrupted, for example, the efficient functioning of the *evaluator*, which might have had to check the proposed ideas and the transcribed material by constantly shifting its focus between L1 and L2 while trying to consider the updated/maintained audience (needs, etc.) simultaneously. Moreover, because the audience-oriented task successfully manipulated and embedded the audience into the

writers may be creating their L2 sentences by translating from L1, while relatively higher EFL writers directly generate their texts in L2 (Wang & Wen, 2002), I conclude that this is the case for the intermediate (lower level) writers. To note, this observation was supported by the literature (e.g., Silva, 1993).

⁷⁸ Having said that, once again, I accept that this study did not aim and systematically assess these processes, this data comes from simple background check of the groups and the future studies should investigate these processes further to reach a clear conclusion.

writing processes, this load's influence on the writing processes⁷⁹ might have been felt by the participants who did not have sufficient EF (or trait PT) resources to deal with this challenge more.

7.1.5 The overall conclusion

This study designed an audience-oriented task and investigated its influence on the writing performance of upper-intermediate novice EFL writers by considering two psychological variables, trait PT and EF.

Regarding the overall writing quality, an audience-oriented writing task that directs the attention of upper-intermediate level novice EFL writers to the audience may be beneficial for increasing the overall writing quality of the participants with a relatively higher EF capability or a relatively higher PT ability. This effect seems to be above and beyond the L2 background of those EFL writers. Other than that, among the variables investigated in this study, the essential contributor is the EF capability of the participants. When it comes to the question of whether trait PT may compensate for the absence of enough EF resources, although the data does not allow us to make a clear conclusion, higher trait PT may partially compensate and approach the performance of these participants to the higher EF group when their attention was explicitly directed to

⁷⁹ I still recognize that it may be important to consider the psychological dimension of the issue. For example, as the students in this study, and the participants in some research (e.g., Fujii, 2012), reported, although it is not ideal using the L1 resources can make EFL writers less stressed. Thinking that stress may be the other reason for loading the limited resources (Schoofs, Preuß, & Wolf, 2008), and the use of L1 in an L2 context can reduce the stress and cognitive load under certain circumstances (Bruen & Kelly, 2017), translating ideas from L1 may ease the system from another source of extra load. Thus, I recognize the strategic importance and benefit of L1 use, particularly in lower proficiency contexts. Moreover, some researchers suggest that the EFL writers may feel the difficulty in orchestrating the cognitive writing activities and revert to L1 to avoid overload, which may result in lower-quality passages (see Weijen et al., 2009). Hence, it seems that more investigations are needed to clarify the overload and writing quality issues for the participants from different proficiency levels.

the audience. Lastly, the alternative task, which included a kind of brainstorming rather than PT, may be more beneficial for those participants who do not have either high EF capability or high trait PT.

Regarding the level of persuasiveness, it seemed that the challenging nature of creating persuasive arguments/passages resulted in an overall floor effect. There was no interaction between the task (audience-oriented experimental versus non-audience-oriented control) and trait PT (high versus low) and there was no main effect of the task, trait PT or EF (high versus low). Still, the analyses revealed a marginally significant interaction between the tasks and EF capability which may signal that the experimental audience-oriented task's success in increasing the level of persuasiveness of the passages may depend on the EF capability. Those novice EFL writers with higher EF capability benefitted from this kind of audience-oriented task probably because they could deal with the overload created by less proceduralized L2 language processes and active audience processing while they were trying to adapt convincing reasons. Lastly, like in the overall writing quality, if the writer does not have enough EF resources to allocate, the task directing the attention of novice EFL writer to the audience might have a detrimental effect.

Thus, the overall conclusion is that even a single writing task may be influential in directing the writer's attention to the audience, it may interact with some individual characteristics such as EF and trait PT and increase the upper-intermediate novice EFL writers' overall writing quality performance. Regarding the level of persuasiveness of the passages, the audience-oriented task's influence may depend on the availability of EF resources. These findings confirm the importance of the task and trait PT in EFL

writing and establish the central role of EF in successfully integrating the task with the other processes.

7.2 Further discussion and future directions

This study revealed the importance of designing a task to direct EF resources, employ trait PT abilities, and increase overall writing quality (and to some extent the level of persuasiveness). However, the influence of the task was restricted to those with either higher EF capability or trait PT tendency. Although the manipulation check revealed that they maintain the audience in their minds, those with low EF capabilities, particularly the low EF – low PT group, did not benefit from this. As mentioned above, one possibility for this observation may be the lack of enough EF resources (or compensating trait PT) to integrate the activated audience with writing processes. This follows the essential underlying suggestion of cognitive models: "Writers are constantly, instant by instant, orchestrating a battery of cognitive processes as they integrate planning, remembering, writing, and rereading" (Flower & Hayes, 1981, p. 387) and the literature which reported that EF resources are required to orchestrate writing by allocating attention most efficiently (Kormos, 2012; Weigle, 2005). Thus, novice EFL writers may not be quite aware of their audience, but this may not be the only problem: they may also have difficulty successfully integrating them into writing processes (e.g., revising) (Kellogg, 2008), particularly if they do not have enough EF capacity (McCutchen, 1996).

As a result of this integration challenge, as Hayes (1980b) suggested, they may be strategically ignoring and showing low audience awareness. Hence, an ideal task

must not only direct the attention and increase audience awareness but also encourage novice writers to integrate the writing processes (e.g., monitoring and revising) by using their available resources at an optimum level⁸⁰. Future studies can design these kinds of tasks, such as asking novice writers to monitor what has been written so far, read it from the audience's perspective, and revise it if needed.

However, it may be essential to emphasize that increasing task complexity and directing attention resources may not always guarantee beneficial results. While Cognition Hypothesis (Robinson, 2007) suggests that the direction of attention resources through task instructions can increase the mental effort (the level of notice) of the writers, can help them to make cognitive distinctions in language processes, and can improve their writing (Robinson, 2007; Robinson, & Gilabert, 2007), the Limited Attentional Capacity model/Trade-off Hypothesis (Skehan, 2009) suggests that WM capacity is limited, and because of this, the direction of the attention through a task can result in a trade-off between different dimensions; that is, there can be an improvement on only one dimension. The latter hypothesis seems to be more explanatory for the results obtained, particularly from the low EF group of this study. As I reported above, the control task (including a kind of simple brainstorming part) could be more beneficial for the low EF – low PT group than the experimental task (including a state PT part), which decreased their performance. In other words, although directing attention helped participants take other people's perspectives (Apperly et al., 2006) and improved the

⁸⁰ This is what skilled writers do. They adapt different strategies for different tasks and manage the weight of their writing processes, for example, by increasing engagement with planning processes (Beauvais, Olive, & Passerault, 2011). Otherwise, particularly if a novice writer's EF capability is not so high, even when they are or become aware of the audience, they may have difficulty successfully integrating the audience into some writing processes; for example, depending on his EF capability, a writer's audience representation might be actively involved for adapting word choices while transcribing, but might not be maintained and involved for reviewing the text (Kellogg, 2008).

overall writing quality of those who have enough cognitive resources, the performance of those with lower EF capabilities and trait PT was worsened. Hence, activating all the critical components in those cognitive writing models (e.g., Hayes, 1996; Kim & Park, 2019) may not always increase the writing performance, and it may be essential to create and investigate alternative (e.g., more straightforward) tasks which can be more beneficial under certain circumstances, for example, for novice EFL writers with relatively lower trait PT and EF capabilities. Future research may consider their study context and focus on these kinds of tasks, as well.

The other variables that can determine the content of these tasks may be the writing proficiency and language proficiency (which may be necessary but not sufficient for writing well) level of writers. These proficiency levels may interact with distinct individual characteristics and various tasks, which can have divergent reflections on the written products. For example, I purposefully focused on novice EFL writers who were still struggling with language processes. Because of that, I could not analytically assess some sophisticated dimensions, such as the employment of counterarguments with their refutations. However, as the dynamic relations hypothesis suggests, the weight of associations between component skills and the written composition (assessed dimensions) may change with the development⁸¹ (Kim & Graham, 2021). For example, transcription processes/skills can determine writing performance at the beginning (Kim & Park, 2019), but discourse oral language can gain more importance with the

⁸¹ Another thing suggested by this hypothesis is that it may be important to consider the appropriateness of the dimension; for example, suppose that the focused dimension is the quality of the ideas. In that case, the ability to use vocabulary and grammar knowledge, the ability to consider audience needs (PT), and the ability to arrange the passage coherently, as well as the ability to transcribe well, must be important, but for example, if the focused dimension is the productivity (the length) of the passage, PT would not have much to do with it (Kim, 2020; Kim & Graham, 2021).

development (Kim & Graham, 2021). With respect to development within an EFL writing context, having automatized low-level linguistic processes (Kormos, 2012) and writing processes (Becker, 2006) and having richer conceptualizations of their audiences (or at least imagining them) (Kellogg, 2008), skilled writers are more aware of the audience expectations, can take their perspective (Weigle, 2005), and adapt the text for the audience accordingly (Carvalho, 2002). In other words, rather than working on their texts at the surface level, they can work on the content and the organization of their writings, considering the needs of their audience (Weigle, 2005). Thus, future studies may recruit participants from, for example, higher language proficiency levels, or even native writers, whose performance are less dependent on fundamental language processes. Doing this, they may focus on more sophisticated audience awareness reflections, such as the employment of counterarguments with their refutations (Ferretti et al., 2000) or higher-level PT within the passages (Cho et al., 2021).

Again, as the cognitive writing models suggested, the background of the participants related to the topic of the task may be essential to consider. In this study, I conducted pilot studies, after which I eliminated a less familiar topic (city projects) and included a more familiar topic (student clubs) for the students. Additionally, some vignettes that included brief but specific information about the student clubs were presented during the task processing stage. Probably because of this, both control ($M = 1.93$) and experimental ($M = 1.75$) groups reported low difficulty levels (on a five points scale) in understanding the expectations of club members. Moreover, ANOVA analysis revealed that a special interest in the clubs or being involved in related activities did not interact with the experimental tasks and differentially influence the overall writing

quality or the level of persuasiveness. As mentioned before, this may be because the topic was relatively familiar to the students, and the presented vignettes introduced some additional background.

Still, because novice writers may be dependent on knowledge background (Bereiter & Scardamalia, 1987; Kim & Park, 2019) and knowledge that is relevant to writing, together with linguistic skills, can be important to overcome the limitations of EF and to write better passages (McCutchen, 2011), future studies may focus on the topic familiarity and systematically investigate its interaction, for example, with an audience-oriented task to see the change in the adapted strategy and the increase in the level of persuasiveness. For instance, other than assessing the topic background by creating questionnaires, the participants may be exposed to an audience-oriented task and to a resource text in which the level of background information is manipulated. Then, for example, the employment of knowledge telling (directly using) versus knowledge transforming (using after adapting) strategies by different kinds of participants (e.g., those who have higher EF) and the level of persuasiveness of the produced passages may be investigated.

Additionally, in this study, based on previous studies (e.g., McCutchen et al., 1994; Olive, 2011), I assumed that the fundamental language processes consumed EF resources of novice EFL writers and did not assess this suggestion directly. Actually, I explored the correlations between EF and fundamental language background (vocabulary and grammar exam scores), but this does not tell much because these tests assess literacy (Kim & Crossley, 2021) which is a part of LTM (Hayes, 1996) Future

studies may design online vocabulary and sentence generation tasks and consider their interaction with EF resources and high-level writing processes.

Related to this, the interaction between L2 proficiency and EF may also deserve to be considered. For instance, EF is associated with grammatical language processing at low proficiency levels (Serafini & Sanz, 2016), but with the increase in proficiency, it becomes associated with sophisticated word use (Vasylets & Marin, 2021) which in turn is associated with writing quality (Lee, Ge, & Chung, 2021). Future studies may include online grammar and vocabulary (e.g., word retrieval) processes (e.g., similar to Kellogg, Turner, Whiteford, & Mertens, 2016) and investigate their possible interaction with the limited cognitive resources, high-level writing processes, and, for example, whether they have a differential effect on the final written products of the EFL writers from different proficiency levels.

Another future direction may be related to the revision processes. This study aimed to assess the influence of an audience-oriented task while writing and, for this purpose, introduced the manipulation before starting to write and asked the students to compose their passages on a paper. However, as reported above, some researchers suggested that the influence of audience awareness may be more observable during the revision stage and introducing the task after writing the first draft, but before the revision (e.g., Midgette et al., 2008; Roen & Willey, 1988; Wang, 2016) may reveal stronger audience awareness effect. For example, taking reader perspective at this stage may help writers to process possible reader interpretations better and encourage them to revise the passages so that they may convey the meaning more clearly (Traxler & Gernsbacher, 1993).

Although I do not think the audience processes would be highly integrated into the other writing processes (planning) when the manipulation was done after writing the first draft, I agree and would expect that giving the task before revision stage would reveal a stronger effect (because more resources would be free). Thus, to investigate the underlying dynamics and the interaction patterns between the writing processes, future studies may focus on the timing of the audience awareness manipulation, consider the psychological variables (EF and trait PT), and compare the differences between giving the task before or after the first draft.

Beyond these, one thing with respect to revision processes may be related to transcribing technology or composing medium, which can also influence writing processes (Hayes, 1996; Hayes & Berninger, 2014). Although it was beyond the scope of this study and I did not directly assess it, revising is a critical component (e.g., Kellogg, 2001), and understanding whether using papers rather than computers influenced it may deserve further discussion. In this study, considering the availability of resources, I collected paper-based handwritten passages composed in a classroom environment.

One of the concerns related to the paper-based data collection may be about its possible restriction in revising the passages while or after writing. Some researchers (e.g., Kobayashi & Rinnert, 2018) suggested that both computer-based and paper-based tools are similar in many cases (e.g., in planning and evaluating), but computers can be beneficial, especially as concordancing tools for making micro level (e.g., formulating and revising) and macro level (e.g., content and organization) revisions easier. However,

this benefit may depend on contextual factors such as the task requirements (e.g., the length of the passage to be written).

Probably because of the contextual factors, the studies reported in the literature revealed different patterns with respect to the influence of the transcription medium on writing performance and processes. For example, some studies indicated that writing by using a keyboard may increase the performance of EFL learners (e.g., Zhu, Mark Shum, Brian Tse, & Liu, 2016) and, in the long run, using Word processors can have some facilitatory effects (see Graham & Perin, 2007). For example, when advanced EFL writers were considered, it was shown that using computers rather than papers as a writing medium increased engagement with higher-level processes, revealed significantly more revising activities, and resulted in better argumentation patterns (Li, 2006). On the other hand, some researchers indicated that compared to computer-based, paper-based writing could ease some EF resources (e.g., Kellogg, 2001) and may increase writing performance (e.g., Breland, Lee, & Muraki, 2005). And particularly, EFL writers with low language proficiency who may not have enough experience using computers to compose their texts may have to deal with additional cognitive demands when they are asked to write using a keyboard (Wolfe & Manalo, 2004).

Seeing that the findings were contradictory and thinking that it could be better to understand the influence of collecting paper-based data in this study's context, I got in touch with four participants and, after sharing the pictures of the task and their texts, asked them three questions. They were whether using a pen-paper made a difference with respect to revision processes, whether writing the same text in a computer would make a difference, and whether using a pen-paper made them think more carefully

before transcribing. A bit surprisingly, probably because the task did not require writing a very long passage, only one participant favored computer over pen-paper and suggested that it would be better because tracking and editing the text (errors) would be easier, she could write faster by using a computer, and she would have more time for reviewing (Participant 625). Other than her, the main reasons for those who favored a pen and paper were: using pen-paper allowed a participant to translate her ideas more comfortably (Participant 403), looking at the screen for a long time might make another participant tired (Participant 911), and using pen-paper made a participant feel psychologically more comfortable, helped her focus on the topic and writing, and editing in the computer would be difficult for (and distract) her (Participant 566). Additionally, a follow-up question revealed that except for Participant 625, who reported it would not make a difference; the participants reported that using a pen-paper might have encouraged them to plan and think more before transcribing the ideas. For example, Participant 403 suggested that using pen-paper allowed her to plan by taking short notes and not to forget her ideas; Participant 911 said that using paper, via taking short notes, allowed her to think more about the topic; and Participant 566 reported that using paper helped her to write by thinking in more detail.

Thus, based on this restricted data and its context, I tend to think that using a paper rather than a computer might have eased some EF resources (e.g., Kellogg, 2001) or at least did not interrupt the revision⁸². On the other hand, the alternative transcription medium (computer) could overload WM resources and interrupt text generation (e.g., difficulty in tracking ideas and the reduced linguistic accuracy) particularly if a student

⁸² I also asked these questions to the students with the same level as the study participants, and my observation is that most of them would prefer writing by using a pen-paper rather than a computer.

is not fluent in typing (McCutchen, 2011)⁸³. Because almost every adult is fluent in handwriting, I assume that writing by using pen-paper were safer and did not create such an extra load.

Lastly, although there is not enough data to evaluate this, and this was not the aim of this study, it may be important to emphasize that concerning audience does not only influence the revision processes. Maybe more critically, it may influence processing the ideas mentally and choosing appropriate words before writing (Phrases, 2002). That is, revising (or evaluating) does not necessarily occur after transcribing. It may be done during the formulation of the ideas (Hayes & Berninger, 2014), which include pre-text/tentative formulation processes (e.g., modifying or simplifying before writing) (Zimmermann, 2000). Exploring this possibility (whether the participants in the experimental group evaluated more before transcribing) can be another entertaining topic for future studies.

7.3 Pedagogical implications

The main results of this study indicated that the experimental task, which explicitly directed the attention of the novice EFL writers to their audience, triggered audience-oriented processes and affected the writing performance of upper-intermediate novice EFL writers by interacting with EF and trait PT. Those with relatively higher EF

⁸³ Still, I recognize the importance of computers, particularly for revision research, but I also want to emphasize that their usefulness may be dependent on the timing of the manipulation and the digital literacy of the participants. For instance, some studies indicated that reading the written passages and then revising them may result in better drafts (Moore & MacArthur, 2012), particularly if the audience's perspective is taken (Traxler & Gernsbacher, 1993). Thinking that working on an already written passage may be easier for some EFL writers, future studies may consider the transcription tool skills and preferences of the participants and follow, for example, their revision patterns on a computer screen to get insight into the interaction of audience manipulation with different kinds of revision processes.

capability or trait PT tendencies benefitted from this audience oriented experimental task, which included an embedded state PT activity. On the other hand, the experimental task worsened the performance of those with neither high EF capability nor trait PT tendencies (low EF - low PT). Actually, they performed better in the control condition, which included a kind of simple brainstorming activity. This shows the importance of the task. If a task is created appropriately, it may help to manage these psychological resources and improve writing performance. Hence, the instructors (and students) must be aware of the importance of a task and the involvement of critical psychological components, such as EF and trait PT.

This and many following implications are in accordance with the process of writing pedagogies⁸⁴, which emphasize awareness and intervention as its core components (Susser, 1994). As Susser (1994) reported, according to this approach, writing is not only getting something clearly in mind and then writing it down. Rather than that, writing is a process of discovery in which the ideas are generated but not necessarily transcribed. Thus, the students and instructors must be aware that even very simple messages include writing processes such as choosing lexical items, considering the audience, and evaluating the format. The second component of this writing as a process approach is intervention. Susser (1994) added that although it is not necessary, an intervention may be done by the process teachers (or by peers, etc.), for example, by introducing brainstorming activities⁸⁵. Importantly, he also emphasized that the teachers

⁸⁴ I recognize the existence and importance of the other, for example, genre/social approaches which emphasize the social context of the writing activity. However, because this study focuses on cognitive writing processes, I primarily focus on process-based (rather than other) suggestions.

⁸⁵ There are studies that revealed different alternatives which may reveal similar easing the sources effects. For example, Kellogg (1998) indicated that although creating drafts does not help, creating a

must consider individual differences/processes and accommodate their tools according to the needs of the students. They must create/use the tasks which may have the potential to help the writers ease or manage the resources accordingly. However, this does not mean always giving tasks that do not create cognitive loads. As some researchers (e.g., Byrnes & Manchon, 2014) suggested, instructors must recognize the importance of having a variety of tasks in their toolboxes and use tasks from a spectrum that encompasses both simple and cognitively challenging ones (beyond the linguistically and cognitively simple ones) to respond to the needs of different learners.

The second implication may be related to the importance of audience processing during writing. From the literature, we know that the audience may be embedded into the mental representations of a writer while processing a writing task (Wang, 2016), and many experimental studies reported that concern for the audience results in some adaptations which may contribute to the quality of the passages (Midgette et al., 2008; Phrases, 2002; Wang, 2016). However, as we also know, although more skilled writers adopt a kind of knowledge-transforming strategy and tend to re-elaborate the retrieved LTM information in order to, for example, clarify new ideas while maintaining the audience in their minds, novice writers tend to adopt a kind of knowledge-telling strategy and do not consider the audience while, for example, planning, organizing or employing the other writing processes (Ling, Choy, & Wei Jhen, 2022).

One way to deal with this problem may be adopting the procedural facilitation method, which employs simplified external aids to proceduralize and integrate the

written or mental outline may contribute to easing the attentional overload and increase the overall quality.

writing processes (Carvalho, 2002). For example, when the writers feel it is needed, they may stop and adapt their passages for their audiences, for example, by asking questions about whether their reader may find something unclear (Carvalho, 2002). Based on the observations of this study, I suggest that introducing an audience-oriented task that can embed a constantly maintained audience into those novice writers' mental processes may also be a proceduralization step. These kinds of tasks may help novice EFL writers' minds be entrenched with, for example, audience processing⁸⁶. And in the long run, they may be proceduralized and integrated better with the other writing processes, such as reviewing. Hence, they may be helpful for those instructors who want to consider the writing processes as they scaffold novice EFL writers by increasing their awareness about the existence of the audience⁸⁷.

Thirdly, as suggested by Cho and Choi (2018), not mentioning the audience in a writing prompt may give the impression that the aim of writing is evaluation rather than real communication with the audience. Our data provided tentative support for this suggestion. Compared to 25 out of 62 in the experimental group, 34 out of 65 participants in the control group reported that their writing would be evaluated within the testing context (e.g., by a teacher) (see Table 23). Moreover, these novice EFL writers perceived the experimental task as less difficult than the control task (see Table 19), and more than half of the participants who perceived main writing as significantly

⁸⁶ These kinds of interventions would also be in accordance with the process writing approach, which involves writing for authentic audiences while encouraging the writers to engage in cycles of planning, translating, and reviewing (e.g., instructing how to add, delete, rearrange the ideas, etc.) (see above and also Graham & Perin, 2007).

⁸⁷ As Susser (1994) emphasized, the breakdown of the writing processes into separate isolated stages could violate the recursive nature of the writing. Thus, it may be important to embed these kinds of processes into recursively created and applied activities which may encourage increasingly complex (multiple) processes.

difficult (22 out of 41) reported that their passages would be evaluated in a traditional testing context. Hence, thinking that the only difference between the tasks was the level of audience focus, although this association should be investigated further by future studies, the teachers may consider the possibility that presenting these kinds of audience-oriented tasks may create a relatively more authentic⁸⁸ (in its perceived evaluation) and psychologically safer writing environment, which may facilitate learning.

The fourth implication is related to the importance of PT. It was previously shown that PT has great general benefits, such as improving intergroup relations by merging the self with the other and encouraging the perspective taker to recognize the differences (Todd & Galinsky, 2014), decreasing in-group favoritism (Galinsky & Moskowitz, 2000), or increasing creativity (Hoever, Van Knippenberg, Van Ginkel, & Barkema, 2012). These kinds of findings can give educators more than enough reasons to embed it into the curriculum. Beyond that, writing with an awareness of the audience requires understanding the needs and perceptions of the reader (Ling, Choy, & Wei Jhen, 2022), and the studies indicated that PT, which is a high-level component associated with audience processing, contributes to writing performance (e.g., Kim & Park, 2019). The present study confirmed these suggestions and revealed that the psychological trait PT could interact with writing processes (via an audience-oriented writing task) and increase general writing performance. This may be another reason to make it part of an EFL curriculum.

⁸⁸ Authenticity may be important because as they respond to a more authentic audience, the students employ more critical communication processes, such as reviewing their ideas according to how the audience perspective might reflect back (Magnifico, 2010).

There may be several ways of doing this. For example, similar to a task set used by Hooper, Erdogan, Keen, Lawton, and McHugh (2015), who adapted from McHugh, Barnes-Holmes, and Barnes-Holmes (2004), it can be done directly by using trials that ask the participants to respond by changing his/her perspective between the reversed “I–YOU, HERE–THERE and NOW–THEN”. To illustrate one of these trials, someone who practices a “I–YOU” reversed trial would answer: “I have a red brick, and you have a green brick. If I was you and you were me, what would you have?” (Hooper et al., 2015, p. 70). Hooper et al. (2015) indicated that practicing trials like this improved the participants’ PT and decreased fundamental attribution errors. It is also possible to embed these PT training processes into the EFL classroom context more indirectly. For example, as one of my professors shared⁸⁹ the instructor may simply present some pictures taken in different contexts and ask the students to speculate what kind of experiences, expectations, etc., the people in the picture have (e.g., whether they have a family, what the person was doing five minutes before, etc.). Again, some structured classroom discussions can be organized to improve trait PT. As Chadwick and Ralston (2010) indicated, compared to unstructured ones, these kinds of class discussions can also increase the level of PT (e.g., more mutual communication), which may become associated with students’ academic performance (higher grades).

The fifth pedagogical implication is related to another observation of this study, the lack of sound reasoning (argumentation) in the persuasive passages. Argumentation is a valuable skill required by many disciplines in university life (e.g., empirical arguments related to data analyses or text-based arguments helpful in interpreting a text

⁸⁹ I thank Professor Senem Yıldız for sharing these kinds of methods in her intellectually entertaining lectures.

from different perspectives, etc.) (Wolfe, 2011). However, many undergraduates cannot employ its fundamental components, such as providing supporting evidence and its association with persuasiveness (Wingate, 2012). Hence, appropriate interventions may be designed to equip those novice writers with skills and strategies for improving their reasoning abilities.

When it comes to how to do this, as Ferretti and Fan (2016) suggested, the capacity limitations (and my-side bias) may be eased by the interventions, which may include dialogic interactions (e.g., Wen et al., 2023) and self-regulatory supports (e.g., Mason & Shriner, 2008). For example (Kuhn & Crowell, 2011) designed a longitudinal study and tasks that encouraged dialogic interactions, which aimed to increase the participants' thinking abilities about the given topics. The intervention included within-group *pre-game* (e.g., evaluating their and others' reasons), the within-group *game* (the same side participants were paired and argued against the other pairs during 25 min.), and within plus between groups *end-game* sessions (the participants made their last preparations for a shot down debate and argued against the other group). Finally, they were debriefed about the argumentation patterns and were asked to write an essay about the topic (see Kuhn & Crowell, 2011, for details). It is suggested that although it may be difficult to control, these kinds of carefully designed dialogic interactions may encourage PT⁹⁰ and improve persuasive writing skills (Ferretti & Fan, 2016). Again, pedagogical support for self-regulatory processes may be useful. For example, MacArthur, Philippakos, and Ianetta (2015) investigated teaching self-regulatory

⁹⁰ It was shown that intervention with tasks similar to these, such as collaborative small-group discussions, may improve PT ability, as well (Wen et al., 2023).

strategies such as the ones for goal setting, progress monitoring, and self-evaluation of writing and indicated that embedding these into a curriculum, compared to the control group, adds 1.5 additional points on a 7-point quality scale. Hence, EFL instructors may consider these methods to contribute to students' reasoning performance.

Last, it is sure that, other than the students and teachers, these findings have some implications for curriculum designers, material developers, and testers. All the implications mentioned above should be considered by those parties, as well. For example, curriculum designers may create syllabi that aim to respond to the needs of EFL learners with different psychological characteristics (e.g., high EF, low PT, etc.). The programs may consider proceduralizing audience and PT processes by integrating them into the learning activities. Material developers may create a range of audience-oriented tasks (including PT training) that encompass simple and cognitively challenging ones. They can design materials that increase the writers' awareness and help them follow some proceduralization steps (e.g., emphasizing the audience's existence while revising, etc.) for progressively entrenching EFL learners' minds with audience perspective. Both curriculum designers and material developers should also consider the challenging nature of argumentation (creating persuasive passages) for EFL learners and plan and develop activities such as dialogic interactions to improve their reasoning ability (see above). Again, testers should be aware that the presence or absence of the audience in the testing materials may create a difference. Additionally, depending on the material they created, the rubrics may include audience-related criteria to give appropriate feedback to EFL learners.

7.4 Contribution

The previous studies indicated the importance of EF (Vasylets & Marin, 2021) and PT (Cho et al., 2021), which covers audience awareness (Kim & Park, 2019). Again, some studies indicated the importance of embedding the audience component into a writing task (e.g., Block & Strachan, 2019; Cho & Choi, 2018). However, to my knowledge, no previous studies considered EF and trait PT and manipulated a writing task that directs EFL writers' attention to actively maintain the audience while writing. This study filled this gap and contributed to the literature by showing that an audience-oriented writing task may interact with EF and or PT and increase the overall writing quality (and, to some extent, the level of persuasiveness) performance of upper-intermediate level novice Turkish EFL writers. Moreover, it indicated that trait PT may possibly compensate for the lack of EF resources when the task requires an explicit orientation to the audience. Besides these, this study confirmed the importance of cognitive resources in composing and supported the models emphasizing EF and PT's contribution.

7.5 Limitations

The study had several limitations. First, I am aware that between-subject designs may have some drawbacks, such as being more conservative and requiring more statistically robust results. More importantly, in this type of design, participant characteristics may be less controllable than within-subject designs. Still, as reported by many researchers (e.g., Charness, Gneezy, & Kuhn, 2012), both designs have their own merits and drawbacks, and the study context must be considered for adopting the more appropriate one. For several reasons, I chose a between rather than a within-subject design for

conducting this experimental study. For example, exposing the participants to experimental or control writing conditions and then asking follow-up questions that aimed, for example, to assess the efficiency of manipulation, would create a carryover effect. Additionally, even counterbalancing the writing tasks might not compensate for the possible confounding, such as experimenter demand effects (Charness et al., 2012).

Still, I took some steps to reduce the drawbacks of the between-subject design. First, I tried to match the groups concerning the most relevant characteristics, such as their gender and proficiency level. Second, I checked the participant's backgrounds via a series of questions (e.g., their interest in student activity clubs and their general writing background) and coded and compared their answers to see if any statistical or non-statistical measures must be taken.

The other criticism related to the between-subject design of the study may be about, similar to some research in the area (Nussbaum, Kardash, & Graham, 2005), collecting the main writing data in a single session after familiarizing the participants with a diagnostic writing task. I had some reasons for doing this. First, it is ideal for collecting main-writing data at several sessions after keeping all possible confounds constant (e.g., the specific content processed in a classroom and the more influential carry over-effect in the experimental groups). However, thinking that the data was collected from 12 classes, this would not be possible. Second, and more importantly, this experimental study was built upon a cognitive approach by putting the limited EF (which processes information within a short temporal window) into its center. Based on this, I wanted to reduce the possible contribution of LTM (carryover effect) and preferred to present the experimental task in a single sitting. Furthermore, thinking of

the experimental nature of the design (the aim was not, for example, to assess proficiency and determine the level of participants or to assess the efficiency of an instruction method over a period of time, etc.), this methodology which many researchers applied may be acceptable.

Still, I tried to increase the number of both experimental and control group participants, and instead of collecting, for example, four main writing passages in four separate sessions from 30 participants, I collected one main writing passage but from 133 participants. Doing this helped to get a more normal distribution and reduced the noise/standard error as much as possible. Additionally, to check if this caused an extra perceived writing difficulty, I collected retrospective reflections on writing and compared the perceived writing difficulty with perceived general writing difficulty. The comparisons revealed that the experimental writings may be perceived even less difficult.

The third concern may be about choosing participant samples from upper-intermediate novice writers who were to follow their studies in different departments. Thinking that they would not automatically employ language processes and the previously stored templates, and the involvement of EF (and trait PT) would be more observable (McCutchen, 2000)⁹¹, I collected the data from novice EFL writers who were not systematically exposed to department-based writings. Other than that, if I had collected the data from the departments, it would have been challenging to control their

⁹¹ This is because the participants who still struggle with basic language (and writing) processes may recruit their WM more actively compared to relatively more expert writers, and the effect of the experimental task could be more observable particularly for novice writers (McCutchen, 2011), and particularly if their language generation processes are not fluent enough (McCutchen, 2000; Schmidt, 1992).

exposure to different types of departmental lessons, and it would be necessary to control the disciplines. However, collecting the data from these preparatory school students who were exposed to similar materials and still struggling with language processes reduced the effect of this possible confounding. Still, I accept that the students who plan to get an education in different areas could have some tendencies, thinking styles, etc. Because of that, I tried to control and distribute them equally across the experimental and control groups. Moreover, I statistically investigated the possible influence of the departments on the main dependent and manipulation check variables, and no significant results were revealed.

Related to this, the fourth concern may be related to the curriculum of the preparatory school. The curriculum did not include separate writing lessons which focused on and systematically taught different genres, tones, registers, etc. The program followed a content-based approach, and except for the proficiency exam, the format of the exams was in accordance with this. Students wrote their answers using the content shared in a previously encountered text without worrying about the traditional essay format. I considered this context and tried to create a main writing task appropriate for the participants' educational background and experimental manipulation. The question was whether experimental manipulation could change the writing pattern of the experimental compared to the control group, and the study did not make any suggestions, for example, about specific genres. Furthermore, as I stated above, collecting the data from such a sample may have some advantages (e.g., observing the involvement of EF) for the purpose of this study.

Still, replicating and extending this study in an EFL environment, where genre-based writing instructions are adapted and students are aware of different genres' requirements, may also be informative. For example, in this study, I focused on the written opinion passages and considered overall writing quality and the level of persuasiveness dimensions. This was appropriate for the purpose of this study, for example, because opinion passages can allow us to observe the reflection of PT (Zrimsek, 2013) and cognitive demand (EF) (Kim & Pae, 2021). However, future studies may consider other genres, such as narratives or even scientific reports to see how keeping the audience in the focus influence different dimensions, such as register, by interacting with trait PT and EF, and the like.

Lastly, I recognize that this study was conducted in a context where novice upper-intermediate Turkish EFL writers started to take advanced-level lessons. The same design could reveal different results in different social environments. For example, the task reflection may depend on the curriculum design variation or educational background. Thus, the results must be generalized with precaution.

APPENDIX A

ETHICAL COMMITTEE APPROVAL FORM

Evrak Tarih ve Sayısı: 27.05.2022-67784

T.C.
BOĞAZİÇİ ÜNİVERSİTESİ
SOSYAL VE BEŞERİ BİLİMLER YÜKSEK LİSANS VE DOKTORA TEZLERİ ETİK İNCELEME
KOMİSYONU
TOPLANTI KARAR TUTANAĞI

Toplantı Sayısı : 32
Toplantı Tarihi : 26.05.2022
Toplantı Saati : 10:00
Toplantı Yeri : Zoom Sanal Toplantı
Bulunanlar : Prof. Dr. Ebru Kaya, Prof. Dr. Feyza Çorapçı, Doç. Dr. Arhan S. Ertan, Doç. Dr. Senem Yıldız,
Dr. Öğr. Üyesi Yasemin Sohtorik İlkmen
Bulunmayanlar :

Turgut Coşkun
Yabancı Diller Eğitimi

Sayın Araştırmacı,

"The Role of Task on the Reflection of Audience Awareness on the Passages Written by Turkish EFL Learners" başlıklı projeniz ile ilgili olarak yaptığımız SBB-EAK 2022/54 sayılı başvuru komisyonumuz tarafından 26 Mayıs 2022 tarihli toplantıda incelenmiş ve uygun bulunmuştur.

Bu karar tüm üyelerin toplantıya çevrimiçi olarak katılımı ve oybirliği ile alınmıştır. COVID-19 önlemleri kapsamında kurul üyelerinden ıslak imza alınmadığı için bu onay mektubu üye ve raportör olarak Yasemin Sohtorik İlkmen tarafından bütün üyeler adına e-imzalanmıştır.

Saygılarımızla, bilgilerinizi rica ederiz.

Dr. Öğr. Üyesi Yasemin
SOHTORİK İLKMEN
ÜYE

e-imzalıdır
Dr. Öğr. Üyesi Yasemin Sohtorik
İlkmen
Öğretim Üyesi
Raportör

SOBETİK 32 26.05.2022

Bu belge, güvenli elektronik imza ile imzalanmıştır.

APPENDIX B
BACKGROUND QUESTIONNAIRE

(Kişisel Bilgi Formu)

Participant Number (*katılımcı numarası*): _____

Gender (please underline) (*Cinsiyet (lütfeñ altını çiziniz)*): Male (*Erkek*) / Female (*Kadın*)

Date of Birth (*Doğum Tarihi*) _____

Major field of study (*Anadalınız*) _____

Your native language (*Anadiliniz*) _____

English Language Knowledge (*İngilizce Dil Bilgisi*)

How old were you when you started learning ENGLISH (underline the range, please)?

(İngilizce öğrenmeye başladığınızda kaç yaşındaydınız (lütfeñ aralığınız altını çiziniz)?)

0-1 2-3 4-5 6-7 8-9 10-11 12-13 14-15 16-17 18-19 20-or-older

What was your English level when you started the university (circle, please)?

(Üniversiteye başladığınızdaki İngilizce seviyeniz neydi? Lütfen yuvarlak içine alın)

Beginner	Pre-intermediate	Intermediate	Upper-Intermediate
Track 1	Track 2	Track 3	Track 4
(A1)	(A2)	(B1)	(B2)

What was your attendance level in the Fall and Spring semester, please express in percentage?

(Güz ve Bahar döneminde derslere katılım oranınız ne kadardı, yüzdelik olarak belirtin lütfeñ?)

Güz (*Fall*): ____ %

Bahar (*Spring*): ____ %

How actively have you tried to learn English in the prep school so far? Please evaluate it on the following 5-point scale (1 = not at all, 5 = completely).

(Şimdiye kadar hazırlık okulunda ne kadar aktif bir şekilde İngilizce öğrenmeye çabaladınız? Lütfen 5 puanlık ölçek üzerinden (1 = hiç öyle değil, 5 = tamamen öyle) değerlendiriniz.)

1	2	3	4	5
Comment (yorum)?				

APPENDIX C

DIAGNOSTIC WRITING TASK

Imagine that you are a government official in charge of researching some topics to help the president shape government policies. Now, the issue is that the government wants to start a hybrid education (half of the lessons will be online, and the other half will be face-to-face) to increase the efficient usage of online resources. But unfortunately, because of the chip crisis, there are not enough appropriate computers in the country.

Thus, the government will choose one or (if possible) two of the following student groups for hybrid education.

In order to choose the student groups which will benefit from the computers and hybrid education more and to distribute the computers accordingly, you are asked to write an argumentative essay and discuss the advantages of online education for:

Primary school students

High school students

University students

Please think about which student group(s) will benefit more from the computers and hybrid/online education.

Then, order them from the one which will benefit the most to the one which will benefit the least. And, please write your passage by providing reasons and giving examples.

You have 40 minutes to finish your writing. Normally, around 250 words are enough, but you can write longer.

If you want, you can use the empty part of this paper to make your preparation.

APPENDIX D
MAIN WRITING TASKS

The Audience Oriented Task (Experimental Group)

Task A Introduction

Imagine that the university management collects some student opinions on how to share 80.000 lira among three different student clubs.

Also, imagine that your opinion will contribute to your university rector's decision. Both the rector and the members of the clubs will read your essay, and then they will organize a meeting to make a final decision to share the money accordingly.

Now, the researcher will distribute some passages.

Please read the messages that these three clubs' representatives sent.

After reading each message, you will be asked to take the perspective of these club members and write whatever comes to your mind. Thus, please try to take the club members' perspective as you read.

Eğer net olmadığını düşündüğünüz bir kısım varsa/olursa mutlaka araştırmacıya sorunuz!

Outdoor Sports Club



Zeynep is the representative of the Outdoor Sports Club. She says, “our club is interested in trekking, mountaineering, and rock climbing. No previous experience is necessary to join our club. The experienced club members will organize both theoretical and practical training sessions so that new club members gain the necessary knowledge and experience. These training sessions include “rock climbing”, “camping”, “mountaineering”, “trekking”, etc. We need approximately 50.000 lira in order to buy some climbing equipment”.

Now, please take the perspective of these club members (their expectations, thoughts, feelings, etc.) and write whatever comes to your mind. You have 3 minutes.

Theatre Club



Merve is the representative of the theatre club, which consists of 45 members. She says, “our club welcomes everyone who is interested in theatre. We organize the training sessions in our club, including stage management, dramas, artistic performances, and exhibitions. We also organize social activities such as parties to increase the connections among club members. In those days, we are in touch with a professional theatre player. We want to invite her and to organize some training sessions. We need 25.000 lira for this activity”.

Now, please take the perspective of these club members (their expectations, thoughts, feelings, etc.) and write whatever comes to your mind. You have 3 minutes.

Music Club



Ali is the representative of the music club, which consists of 6 members. He says, “our club wants to keep the music alive at the university because we believe this will make our university a more peaceful place. In general, our club members come together to talk with each other about the newly produced music, etc. At the moment, we do not have training sessions for the students, but we plan to start some free piano lessons. We need 70.000 lira to buy an acoustic piano”.

Now, please take the perspective of these club members (their expectations, thoughts, feelings, etc.) and write whatever comes to your mind. You have 3 minutes.

In Summary, the rector of your university has received three messages from these three club representatives.

The Outdoor Sports Club members need 50.000 lira to buy climbing equipment and to organize training sessions.

The Theatre Club members need 25.000 lira to invite a professional theatre player and to organize training sessions.

The Music Club members need 70.000 lira to buy an acoustic piano and to organize training sessions.

But the problem is that there is only 80.000 lira which these clubs can use.

The Writing Prompt

Please order these needs in importance and decide how much of the money you would allocate for each club. Write a passage to convince people by providing reasons and giving examples.

As mentioned above, your opinion will contribute to your university rector's decision. Both the rector and the members of the clubs will read your essay, and then they will organize a meeting to make a final decision to share the money accordingly.

You have 40 minutes to finish your writing. Normally, around 250 words are enough, but you can write longer.

If you want, you can use the empty part of this paper to make your preparation.

Participant Code: _____

This image shows a single sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.

Participant Code: _____

[illegible]

The Non-Audience-Oriented Task (Control Group)

Task B Introduction

Imagine that the university management collects some student opinions on how to share 80.000 lira among three different student clubs.

Also, imagine that your opinion will contribute to the decision of management. Both your essay and the other student essays will be considered, the calculations will be made, and the money will be shared accordingly.

Now, the researcher will distribute some passages.

Please read the details which give basic information about these three clubs.

After reading each message, you will be asked to write whatever comes to your mind.

Eğer net olmadığını düşündüğünüz bir kısım varsa/olursa mutlaka araştırmacıya sorunuz!

Outdoor Sports Club



The outdoor sports club is one of the active clubs of the school. The main activities of this club are trekking, mountaineering, and rock climbing. No previous experience is necessary to join the club. There are both theoretical and practical training sessions in the club so that newcomers gain the necessary knowledge and experience. These training sessions include “rock climbing”, “camping”, “mountaineering”, “trekking”, etc. The club needs approximately 50.000 lira in order to buy some climbing equipment.

Now, please just write whatever comes to your mind. You have 3 minutes.

Theatre Club



The theatre club is one of the art-oriented clubs of the school. In the club, many training activities are organized. These activity sessions include stage management, dramas, artistic performances, and exhibitions. The club also organizes social activities such as parties to increase its reputation. In those days, the club plans to organize professional theatre seminars about the analysis of the most well-known plays. One of these plays will be chosen and acted on the stage. The club needs 25.000 lira for this activity.

Now, please just write whatever comes to your mind. You have 3 minutes.

Music Club



The music club is one of the important clubs in the school. The club aims to keep the music alive at the university in order to make it a more peaceful place. In general, club activities include sharing knowledge about the newly produced music, etc. At the moment, there are no training sessions, but the club plans to start some free piano lessons. The club needs 70.000 lira to buy an acoustic piano.

Now, please just write whatever comes to your mind. You have 3 minutes.

In summary, the management of your university has received three messages from these three clubs.

The Outdoor Sports Club needs 50.000 lira to buy climbing equipment and to organize training sessions.

The Theatre Club needs 25.000 lira to invite a professional theatre player and to organize training sessions.

The Music Club needs 70.000 lira to buy an acoustic piano and to organize training sessions.

But the problem is that there is only 80.000 lira which these clubs can use.

The Writing Prompt

Please order these needs in importance and decide how much of the money you would allocate for each club. Write a convincing passage by providing reasons and giving examples.

As mentioned above, your opinion will contribute to the decision of management. Both your essay and the other student essays will be considered, the calculations will be made, and the money will be shared accordingly.

You have 40 minutes to finish your writing. Normally, around 250 words are enough, but you can write longer.

If you want, you can use the empty part of this paper to make your preparation.

Participant Code: _____

[illegible]

Participant Code: _____

[illegible]

APPENDIX E

THE IMAGES USED IN THE EXPERIMENT

Experimental Task

Outdoor Sports Club (Yaşarlı, 2014)



Fig. E1 The outdoor sports club visual used for audience-oriented experimental task

Yaşarlı dağcılar 7. yıllarını Erciyes'in zirvesinde kutladı. (2014, June 10). Yaşar Üniversitesi Haber Portalı. <https://haber.yasar.edu.tr/genel/yasarli-dagcilar-7-yillarini-erciyesin-zirvesinde-kutladi.html>

Theatre Club (“Yastık Adam,” 2019)



Fig. E2 The theatre club visual used for audience-oriented experimental task

Yastık Adam prömiyer yaptı. (2019, May 03). Bursa Haber. <https://www.bursahaber.com/eskisehir/yastik-adam-promiye-yapti-h1560381.html>

Music Club (Ajandakolik, n.d)



Fig. E3 The music club visual used for audience-oriented experimental task

Ajandakolik. (n.d). *Babylon'dan üniversiteli müzik gruplarına destek: Kampüs Konserleri*. Retrieved April 21, 2022, from <https://ajandakolik.wordpress.com/2018/11/26/babylondan-universiteli-muzik-gruplarina-destek-kampus-konserleri/amp/>

Control Task

Outdoor Sports Club (Harmancı, 2020)



Fig. E4 The outdoor sports club visual used for non-audience-oriented control task

Harmancı, S. (2020, June 21). *Terörden temizlenen Cilo Dağları ve Sat Buzul Gölleri milli park olma yolunda*. Anadolu Ajansı. <https://www.aa.com.tr/tr/turkiye/terorden-temizlenen-cilo-daglari-ve-sat-buzul-golleri-milli-park-olma-yolunda-/1884570>

Theatre Club (Kızıldağ, 2020)



Fig. E5 The theatre club visual used for non-audience-oriented control task

Kızıldağ, S. (2020, June 08). 'Tiyatro kenti' Eskişehir'de Kovid-19 sürecinde kapanan perdeler açılmayı bekliyor. *Anadolu Ajansı*. <https://www.aa.com.tr/tr/kultur-sanat/tyatro-kenti-eskisehirde-kovid-19-surecinde-kapanan-perdeler-acilmayi-bekliyor/1868996>

Music Club (Özyeğin, n.d.)



Fig. E6 The music club visual used for non-audience-oriented control task

Özyeğin Üniversitesi Ses Kayıt Studyosu. (n.d.). Medya – Studyo. Retrieved April 21, 2022, from <https://www.mea-akustik.com/ozyegin-universitesi-ses-kayit-studyosu.html>

APPENDIX F

RETROSPECTIVE QUESTIONNAIRE

(Oturum Sonrası Geriye Dönük Değerlendirme)

Bu oturumda son olarak bazı sorulara cevap vermenizi rica edeceğiz. Diğer kısımlarda olduğu gibi bu kısımdaki soruların da dikkatle okunması ve cevaplanması araştırmamız için oldukça önemlidir. Katkınız için tekrar teşekkür ederiz.

(Finally, in this session, we will ask you to answer some questions. As in other parts, it is very important for our research to carefully read and answer the questions in this part. Thanks again for your contribution.)

A - Okuyucu Bakış Açısı (Kulüpler) *(Audience Perspective Taking (Clubs))*

Aşağıdaki ifadelere ne derece katıldığınızı 5 puanlık ölçek üzerinden (1 = hiç öyle değil, 5 = tamamen öyle) değerlendiriniz.

(Evaluate the extent to which you agree with the following statements on a 5-point scale (1 = not at all, 5 = completely so))

1) Mesajlarını okuduğumda, açık hava sporları (outdoor sports) kulübü üyelerinin beklentilerini onların bakış açısından anlamaya çalıştım. <i>(When I read their messages, I tried to understand the expectations of outdoor sports club members from their point of view)</i>				
1	2	3	4	5

2) Mesajlarını okuduğumda, açık hava sporları (outdoor sports) kulübü üyelerinin beklentilerini onların bakış açısından anlamakta zorlandım. <i>(When I read their messages, I had a hard time understanding the expectations of outdoor sports club members from their point of view)</i>				
1	2	3	4	5

3) Mesajlarını okuduğumda, tiyatro (theatre) kulübü üyelerinin beklentilerini onların bakış açısından anlamaya çalıştım. <i>(When I read their messages, I tried to understand the expectations of theatre club members from their point of view)</i>				
1	2	3	4	5

4) Mesajlarını okuduğumda, tiyatro (theatre) kulübü üyelerinin beklentilerini onların bakış açısından anlamakta zorlandım. <i>(When I read their messages, I had a hard time understanding the expectations of theatre club members from their point of view)</i>				
1	2	3	4	5

5) Mesajlarını okuduğumda, müzik (music) kulübü üyelerinin beklentilerini onların bakış açısından anlamaya çalıştım. <i>(When I read their messages, I tried to understand the expectations of music club members from their point of view)</i>				
1	2	3	4	5

6) Mesajlarını okuduğumda, müzik (music) kulübü üyelerinin beklentilerini onların bakış açısından anlamakta zorlandım. <i>(When I read their messages, I had a hard time understanding the expectations of music club members from their point of view)</i>				
1	2	3	4	5

Yazarken okuyucularınızın (rektör ve kulüp üyeleri) bakış açısını dikkate aldınız mı? Aldıysanız, bu durum yazınız üzerinde etkili oldu mu? Nasıl? Detaylandırır mısınız?

(Did you consider the audience (rector and club members) perspective while writing? If you did, did it influence your writing? How? Can you give details?)

B - Yazma Aktivitesi (*Writing Activity*)

Yazma sorusunun anlaşılması kolay mıydı? Detay verebilir misiniz?

(Was the writing question easy to understand? Can you give details?)

Metni yazarken zorlandınız mı? Detay verebilir misiniz?

(Did you experience difficulty while writing the text? Can you give details?)

Metni yazdıktan sonra gözden geçirme şansı bulabildiniz mi? Detay verebilir misiniz?

(Did you have a chance to review the text after writing it? Can you give details?)

C - Değerlendirme Tahmini (*Prediction for Evaluation*)

Sizce yazdığınız bu metin nasıl değerlendirilecek? Detay verebilir misiniz?

(How do you think this text you wrote will be evaluated? Can you give details?)

D - Yazma İçeriğine İlgi (*Interest in Writing Content*)

Müzik, tiyatro veya açık hava sporlarına özel bir ilgi duyuyor musunuz? Detay verebilir misiniz?

(Do you have a special interest in music, theater, or outdoor sports? Can you give details?)

Müzik, tiyatro veya açık hava sporu kuluplerine üye misiniz? Değilseniz olmayı düşünüyor musunuz? Detay verebilir misiniz?

(Are you a member of music, theater, or outdoor sports clubs? If you are not, do you plan to be? Can you give details?)

E - Genel Yazma Deneyimi *(General Writing Experience)*

Genel olarak İngilizce yazma performansınızı değerlendirir misiniz?

(Can you please evaluate your English writing performance in general?)

Genel olarak Türkçe yazma performansınızı değerlendirir misiniz?

(Can you please evaluate your Turkish writing performance in general?)

Yazarken genelde Türkçe mi yoksa İngilizce mi düşünüyorsunuz? Detay verebilir misiniz?

(Do you usually think in Turkish or in English while writing? Can you give details?)

F- Yazma Aktiviteleri (*Writing Activities*)

Derslerde veya ders dışında ne tür yazma aktiviteleri yapıyorsunuz? Detay verebilir misiniz?

(What kind of writing activities do you do in or out of class? Can you give details?)

G – Yazma Eğitimi Geçmişi (*Writing Background*)

İlkokul ve lise dahil olmak üzere üniversite öncesinde İngilizce yazma dersleri aldınız mı? Aldıysanız ders içeriği, materyaller, ödevler, geribildirim vb. ile ilgili detay verebilir misiniz?

(Did you take English writing courses before coming to the university, including primary and high school? If you did, can you give details about the course content, materials, assignments, feedback, etc.?)

İlkokul ve lise dahil olmak üzere üniversite öncesinde Türkçe yazma dersleri aldınız mı? Aldıysanız ders içeriği, materyaller, ödevler, geribildirim vb. ile ilgili detay verebilir misiniz?

(Did you take Turkish writing courses before coming to the university, including primary and high school? If you did, can you give details about the course content, materials, assignments, feedback, etc.?)

APPENDIX G

PERSPECTIVE TAKING SCALE (Davis, 1980)

Please indicate to what degree you agree with these statements on a 5-point scale (1 = not at all, 5 = completely).

1) I sometimes find it difficult to see things from other people's points of view (R).

<i>Not at all</i>				<i>Completely</i>
1	2	3	4	5

2. I try to look at everybody's side of a disagreement before I make a decision.

<i>Not at all</i>				<i>Completely</i>
1	2	3	4	5

3. I sometimes try to understand people better by imagining how things look from their perspective.

<i>Not at all</i>				<i>Completely</i>
1	2	3	4	5

4. If I'm sure I'm right about something, I don't waste much time listening to people's arguments. (R)

<i>Not at all</i>				<i>Completely</i>
1	2	3	4	5

5. I believe that there are two sides to every question and try to look at them both.

<i>Not at all</i>				<i>Completely</i>
1	2	3	4	5

6. When I'm upset at a person, I usually try to "put myself in his shoes" for a while.

<i>Not at all</i>				<i>Completely</i>
1	2	3	4	5

7. Before criticizing a person, I try to imagine how I would feel if I were in their place.

<i>Not at all</i>				<i>Completely</i>
1	2	3	4	5

APPENDIX H

WRITING ASSESSMENT (OVERALL WRITING QUALITY) SCALE

(Jacobs et al., 1981)

ESL COMPOSITION PROFILE				
STUDENT		DATE	TOPIC	
SCORE	LEVEL	CRITERIA		COMMENTS
CONTENT	30-27	EXCELLENT TO VERY GOOD: knowledgeable • substantive • thorough development of thesis • relevant to assigned topic		
	26-22	GOOD TO AVERAGE: some knowledge of subject • adequate range • limited development of thesis • mostly relevant to topic, but lacks detail		
	21-17	FAIR TO POOR: limited knowledge of subject • little substance • inadequate development of topic		
	16-13	VERY POOR: does not show knowledge of subject • non-substantive • not pertinent • OR not enough to evaluate		
ORGANIZATION	20-18	EXCELLENT TO VERY GOOD: fluent expression • ideas clearly stated/ supported • succinct • well-organized • logical sequencing • cohesive		
	17-14	GOOD TO AVERAGE: somewhat choppy • loosely organized but main ideas stand out • limited support • logical but incomplete sequencing		
	13-10	FAIR TO POOR: non-fluent • ideas confused or disconnected • lacks logical sequencing and development		
	9-7	VERY POOR: does not communicate • no organization • OR not enough to evaluate		
VOCABULARY	20-18	EXCELLENT TO VERY GOOD: sophisticated range • effective word/ idiom choice and usage • word form mastery • appropriate register		
	17-14	GOOD TO AVERAGE: adequate range • occasional errors of word/idiom form, choice, usage <i>but meaning not obscured</i>		
	13-10	FAIR TO POOR: limited range • frequent errors of word/idiom form, choice, usage • <i>meaning confused or obscured</i>		
	9-7	VERY POOR: essentially translation • little knowledge of English vocabulary, idioms, word form • OR not enough to evaluate		
LANGUAGE USE	25-22	EXCELLENT TO VERY GOOD: effective complex constructions • few errors of agreement, tense, number, word order/function, articles, pronouns, prepositions		
	21-18	GOOD TO AVERAGE: effective but simple constructions • minor problems in complex constructions • several errors of agreement, tense, number, word order/function, articles, pronouns, prepositions <i>but meaning seldom obscured</i>		
	17-11	FAIR TO POOR: major problems in simple/complex constructions • frequent errors of negation, agreement, tense, number, word order/ function, articles, pronouns, prepositions and/or fragments, run-ons, deletions • <i>meaning confused or obscured</i>		
	10-5	VERY POOR: virtually no mastery of sentence construction rules • dominated by errors • does not communicate • OR not enough to evaluate		
MECHANICS	5	EXCELLENT TO VERY GOOD: demonstrates mastery of conventions • few errors of spelling, punctuation, capitalization, paragraphing		
	4	GOOD TO AVERAGE: occasional errors of spelling, punctuation, capitalization, paragraphing <i>but meaning not obscured</i>		
	3	FAIR TO POOR: frequent errors of spelling, punctuation, capitalization, paragraphing • poor handwriting • <i>meaning confused or obscured</i>		
	2	VERY POOR: no mastery of conventions • dominated by errors of spelling, punctuation, capitalization, paragraphing • handwriting illegible • OR not enough to evaluate		
TOTAL SCORE		READER	COMMENTS	

Taken from Weigle (2002, p. 116).

APPENDIX I

THE LEVEL OF PERSUASIVENESS RUBRIC

(Based mostly on Ferretti, MacArthur, and Dowdy, 2000; and Ong, 2013)

0 The task is not answered appropriately. The passage responds to the topic but does not provide the writer's own position for allocating money to any of the clubs.

1 The passage is not convincing at all. The writer's position about the allocation of money for at least one club is stated. There is some attempt to support the position(s) by providing data (reasons, examples, anecdotes, suggestions, etc.), but most of them are not related and acceptable.

2 The passage is not convincing. The writer's position about allocating money for at least two clubs is stated. The positions are supported by weakly relevant and acceptable data (reasons, examples, anecdotes, suggestions, etc.). The supporting data is not elaborated with details. The passage may contain some inconsistencies.

3 The passage is somewhat convincing. The writer's position about allocating money for at least two clubs is stated. The positions are supported by partially relevant and acceptable data (reasons, examples, anecdotes, suggestions, etc.). The supporting data is somewhat elaborated. The passage may contain a few inconsistencies, if any.

4 The passage is almost convincing. The writer's position about allocating money for all three clubs is stated. The positions are supported by moderately relevant and acceptable data (reasons, examples, anecdotes, suggestions, etc.), and the data is elaborated to some extent. The passage is free of inconsistencies (e.g., in allocating money).

The counter position for the disagreed club(s) (e.g., the money they wanted) may be considered, but no appropriate rebuttal (refutations and/or alternative views) may be proposed.

5 The passage is convincing. The writer's position about allocating money for all three clubs is stated. The positions are supported by quite relevant and acceptable data. The supporting data is elaborated. The passage is free of inconsistencies (e.g., in allocating money) or unnecessary repetitions.

The counter position for the disagreed club(s) (e.g., the money they wanted) may be considered, and their rebuttal(s) (refutations and/or alternative views) may be proposed, they may not be so efficient, though.

6 The passage is quite convincing. The writer's position about allocating money for all three clubs is clearly stated. The positions are supported by highly relevant and acceptable data, and the supporting data are elaborated well. No inconsistencies or unnecessary repetitions are detected.

The counter position for the disagreed club(s) (e.g., the money they wanted) may be considered, and their rebuttal(s) (refutations and/or alternative views) may be proposed. The refutation may be efficient to some extent.

7 The passage is very convincing. The writer's position about allocating money for all three clubs is clearly stated. The positions are supported by perfectly relevant and acceptable data, and the data is explained and elaborated very well. No inconsistencies or unnecessary repetitions are detected.

The counter position for the disagreed club(s) (e.g., the money they wanted) may be considered, and their rebuttal(s) (refutations and/or alternative views) may be proposed. The refutation may be quite efficient.

The Definition of Some Terms Used in the Rubric

Position: It can refer to a writer's position for the argued topic (Majidi et al., 2021) and can be defined as an opinion on the main question (Nussbaum & Kardash, 2005) which is about a controversial topic or problem (Qin & Karabacak, 2010) or as a standpoint which refers to "the writer's opinion on the central issue of the topic" (p. 105) (Wang, 2016).

Counter Position: It can be defined as a criticism or objection that has the potential to challenge a writer's claim (Chase, 2011; Qin & Karabacak, 2010), why his/her "position might not be true or advisable" (p. 157) (Nussbaum & Kardash, 2005).

Rebuttal: It can be thought of as response to counterarguments or "a claim that refutes a counterclaim by demonstrating that the counterargument (a) is invalid, (b) lacks as much force or correctness as the original argument, or (c) rests on a false assumption" (Nussbaum & Kardash, 2005, p. 160). That is, they can be addressed as the conditions which have the potential to defeat (weaken or refute) the alternative claim (Majidi et al., 2021) and to strengthen the writer's claim (Chase, 2011). However, it must be noted that the restatement of a previously stated claim would not be considered a rebuttal (Nussbaum & Kardash, 2005).

Supporting Data: It can simply be defined as evidence that is advanced to support a position/claim (Majidi et al., 2021; Nussbaum & Kardash, 2005) or as the reason (Wang, 2016) whose function is to create a foundation for a position (Toulmin, 2003). They include facts, logical explanations, research data, suppositions, anecdotes, expert opinions, definitions, analogies, etc. (Qin & Karabacak, 2010; Stapleton & Wu, 2015; Abdollahzadeh, Farsani, & Beikmohammadi, 2017). The support for the position can be further elaborated by the data dependent on each other or by the data independent from each other (Zhang, 2019), but ideally, they must be highly relevant to the position.

Important Notes

- 1) The positions for the clubs can be stated in a single or separate statement. For example, participants can say that they would allocate 20.000 lira to Outdoor Sports, 25.000 lira to Theatre, and 35.000 lira to Music Club in a single sentence. Alternatively, they can first say that they would allocate 20.000 lira to Outdoor Sports and provide support for their position, then continue with the Theatre Club, and so on. Both approaches are acceptable.
- 2) Additionally, sometimes although the position is stated, the exact amount of money may not be stated explicitly. For example, a writer may say that “The school must meet the needs of theatre and Outdoor sports club and give the rest of the money to Music Club”. This would be a statement that considered all the positions because it means that “50.000 lira must be allocated to Outdoor Sports and 25.000 lira to Theatre Club, but only 5.000 lira must be allocated to Music Club.
- 3) When it comes to the elaboration of the support provided by the data, it is related not only to the number of independent supporting data but also to the details given for supporting data. A position may be supported mainly by a single data, but it may be detailed so well that the support may be considered as very well-elaborated one. Alternatively, a position may also be supported by some relatively independent (from each other) data. In other words, support for the position can be elaborated by the data dependent on each other or by the data relatively independent from each other (Zhang, 2019). In both cases, the main concern must be the level of relevance and acceptability of the support.
- 4) Another thing related to the support is that sometimes the writers may suggest alternative solutions such as organizing festivals, buying second-hand equipment, donating, etc. Depending on the context, these can be considered as the supporting data for taking a position(s), for example, not allocating money to the music club. In those cases, their relevance and acceptability (plausibility) can be considered, and the passage may be evaluated accordingly.
- 5) Considering a counter position and rebutting is not the necessary requirement of the task, but if a paper employs them, its grade may be increased depending on the level of elaboration of these components, etc., and this is reflected on the rubric. For example, the evaluator may not be so sure whether “the data is quite relevant and acceptable” and whether “the supporting data is elaborated,” but if he/she sees that the writer considered the counter position with its rebuttal, he/she can evaluate the paper as level 5. This is not only because considering the opposite side is important but also because a counter position with a rebuttal can be considered as additional data which supports a position.
- 6) It must also be kept in mind that counter position and rebuttal may be provided only when there is a disagreement between the writer and the club members. For

example, a writer may think that giving 25.000 lira to Theatre Club, and 50.000 lira to the Outdoor Sports Club is okay but giving 80.000 lira to the music club is not. In this case, there could be only one disagreement between the writer and a group of readers, and the counter position may be provided.

- 7) It may also be important to note that not all the descriptors need to apply to score a paper (Jacobs et al., 1981). Sometimes, the criteria from two different levels can apply. In those cases, the evaluator should consider the combination of criteria and choose the most appropriate persuasiveness level.
- 8) Last but not least, the raters must also be aware that the main concern of this evaluation is not language use or rhetorical pattern. They must be careful about the halo effect (being influenced by grammar structure, etc.) (Jacobs et al., 1981), and attend to the content with respect to its persuasiveness.

An additional important note: As mentioned in the methodology section, I chose 16 anchoring papers and shared them with the second evaluator at the third step (see the methodology, Main Experimental Writing Measurements part above). We clarified the levels by evaluating these sample papers separately and discussing reasons for giving them a specific score. See below for some sample papers.

Some Sample Anchoring Papers

Sample 1

Level 0: The task is not answered appropriately. The passage responds to the topic but does not provide the writer's own position for allocating money to any of the clubs.

Main-Writing, Evaluation Code: 898	#Words:192
There are many club for make activities. These are may be music, outdoor, theatre etc. Actually this sports are so it can be funny or exciting. But sometimes club activities generally it is so expensive. Therefore, everyone is cannot easily join. For instance, somebody wanted to join music club. But his not money enoough. So he must more earn money or he can not join this club. In my opinion, this variety clubs it can be more cheaper because each person is want to gain experience and be happy. Apart from many people are working this clubs but almost their have not got client. So may be many clubs are closed. I think this problem is should solve. In addition, this clubs destination to keep the alive at the universit in order to make it a more peaceful world. In those days, clubs are a few arrangements organize for people. Nevertheless people cannot be want because of expensive. I agree this people but this is not normal because it should be for public. According to me money is not everything. I think one of the most important is receive funny from life.	

The Evaluation Comment: *This passage responded to the topic but did not state the writer's position for allocating the money; thus, it got a zero.*

Sample 2

Level 2: The passage is not convincing. The writer's position about allocating money for at least two clubs is stated. The positions are supported by weakly relevant and acceptable data (reasons, examples, anecdotes, suggestions, etc.). The supporting data is not elaborated with details. The passage may contain some inconsistencies.

Main-Writing, Evaluation Code: 708	#Words:204
I think we give Theatre clubs members and outdoor sports club members because money is limited just 80.000 liras. Theatre club needs 25.000 liras and Outdoor sports clubs needs 50000 liras. We can't give the money music club because if we give money to us we don't give money another clubs. We can do something. For example, Theatre clubs, we can give 25000 liras because we can some activities out school. This situation it can be useful for school. Because we will gain some money this activities. in theatree clubs too. We will use for give the money to music clubs. this activities with the gains. Just we need to wait to give the money to music clubs and they need to wait because we dont have to enough money. I hope they understand us. If we give all money the music clubs members wa cant gain too much money. I think we dont have too much opurtonities. therefore we will give the money music and sports clubs. This why, They need to wait for earn the money. If they wait some we will give money to music clubs. I hope music clubs members understand this situation. Because we cant give money all clubs.	

The Evaluation Comment: *This passage responded to the topic, considered the clubs, and allocated the money to two clubs (theatre and outdoor sports). Still, it did not mention what to do with the rest of the money (5.000 lira), and the supports for giving the money to the chosen theatre and outdoor sports were weak. The writer only repeated the problem (there was not enough money) and tried to provide support for giving 25.000 lira to the theatre club (we can do some activities out of school). There was almost no support (apart from not having enough money for three clubs) for the position taken for the outdoor sports club. Beyond these, the most import reason for evaluating this passage as level 2 was that the supporting data (particularly for outdoor sports) were not elaborated. Hence, the overall persuasiveness of the passage is closer to level 2 than the other levels.*

Sample 3

Level 4: The passage is almost convincing. The writer's position about allocating money for all three clubs is stated. The positions are supported by moderately relevant and acceptable data (reasons, examples, anecdotes, suggestions, etc.), and the data is elaborated to some extent. The passage is free of inconsistencies (e.g., in allocating money).

The counter position for the disagreed club(s) (e.g., the money they wanted) may be considered, but no appropriate rebuttal (refutations and/or alternative views) may be proposed.

Main-Writing, Evaluation Code: 443	#Words:215
According to clubs demand total money is not enough. So that the money should divide like; for outdoor sports club 27.000 TL, for theatre club 25.000 TL, for music club 28.000 TL. Outdoor sports club need money to buy equipment and to organize sessions. But, in my opinion we can solve this problem more cheaper. The management of outdoor sports club can offer their members to buy their own equipment. Actually it makes sense because that equipment will be used buy members and also if they accept this offer, 27.000 TL is enough for organizing sessions. If the members do not like this idea, the management of outdoor sports club can offer some companies to give the equipment for free and sign a contract or say that the club will advertise them in university and in the sessions. Theatre club need 25.000 TL totally and the amount of money which will be given by university management is enough for club demand so that they can organize this organisation easily. The music club need 70.000 TL to buy piano. We are offering that they can watch more cheaper pianos or rent to save time. In the future they will organize some parties or organization to sell ticket. And then they can achieve such amount of money.	

The Evaluation Comment: *This passage clearly stated the writer's position for all three clubs (e.g., allocating 28.000 lira to the music club) and supported them with moderately relevant data (e.g., they can buy cheaper pianos ...), which were elaborated to some extent (e.g., alternatively it can be rent ... a party may be organized). There is no inconsistency in the allocation of the money. Beyond these, although no rebuttal was proposed, the counter position for one of the disagreed (music) clubs (for the allocation of the money) was considered (they need 70.000 lira). Mainly for these reasons, the overall persuasiveness of the passage is closer to level 4 than the other levels.*

Sample 4

Level 6: The passage is quite convincing. The writer's position about allocating money for all three clubs is clearly stated. The positions are supported by highly relevant and acceptable data, and the supporting data are elaborated well. No inconsistencies or unnecessary repetitions are detected.

The counter position for the disagreed club(s) (e.g., the money they wanted) may be considered, and their rebuttal(s) (refutations and/or alternative views) may be proposed. The refutation may be efficient to some extent.

Main-Writing, Evaluation Code: 624	#Words:241
I think, we giving money for outdoor sports club So, 50.000 TL for Theatre Club 25.000 TL and for Music club 5.000 TL. Music club members have a good achievement. They want to making more peaceful place. I agree with that guy. Music has a good effect on human body. But I think, They don't need a acoustic piano. Because, piano is so expensive equipment. If club members want to giving lessons, they will use self equipment. For example, someone using guitar, someone using druum. If any equipment brakes or anything, they will use 5.000 TL for the fixing equipment. Maybe they found amore cheaper piano than acoustic piano. For outdoor sports club, They want to 50.000 TL to climbing equipment. I think, we should giving this money. Because climbing has a high risks and you must be use best equipment. For example, If you don't using best equipment, you will injured or maybe die. Also, new equipment has making much easier to learning process. It's so good for the new club members. For theatre club, They want to 25.000 TL to invite professional player and organize some training sessions. I think, we should giving this money, Because this club they have a lot of members. and theatre effect on University social life. Any professional theatre player can learn something club members and training can also learn something other University members. In the same, this organize is cheaper than other activities.	

The Evaluation Comment: *This passage considered all the clubs and clearly stated its position for allocating the money. The positions (e.g., 50.000 lira should be given to an outdoor sports club) were supported by highly relevant data (e.g., climbing has high risks and the best equipment is needed). The supporting data were elaborated (e.g., if you do not use the best equipment, you will be injured ... new equipment can make learning easier). Other than that, there were not inconsistencies or repetitions.*

Additionally, the writer considered the counter position when he does not agree with giving the money to the music club (music club members have a good achievement ... I agree with that guy ...) and, although it is not so strong, a rebuttal with refutation (but I think they do not need an acoustic piano ... they can use a guitar) was proposed. Mainly for these reasons, the overall persuasiveness of the passage is closer to level 6 than the other levels.

APPENDIX J

POST WRITING (BACKGROUND) QUESTIONS CODEGUIDE

Question 1

İlkokul ve lise dahil olmak üzere üniversite öncesinde İngilizce yazma dersleri aldınız mı? Aldıysanız ders içeriği, materyaller, ödevler, geribildirim vb. ile ilgili detay verebilir misiniz?

(Did you take English writing courses before coming to the university, including primary and high school? If you did, can you give details about the course content, materials, assignments, feedback, etc.?)

The answers to this question were coded into three different categories (Table J1).

Table J1. The Categories and Samples for Taking English Writing Lessons Before

Code	Category	Sample
1	Yes, I took writing lessons	Evet aldım, ilkokul ve lisede her hafta sıklıkla İngilizce dersleri görüyordum. Her hafta yazı ödevlerimiz oluyordu. <i>(Yes, I did, I was taking English lessons every week in primary and high school. We had writing assignments every week.)</i>
2	No, I did not take writing lessons	Daha önce almadım. <i>(I have not taken it before.)</i>
3	Other (Not clear)	Okul harici hiç fazladan İngilizce eğitim almadım. <i>(I did not receive any extra English education outside of school.)</i>

Note: In this question, the most important differentiation may be between code (1) and code (3). If a student says something like “Lisede almıştım ama hatırlamıyorum. Burada da öyle üzerine düştüğüm bir şey değil.” *(I took them in high school, but I don't remember. It's not something I care a lot here either.)*, we can code it as (1), we know that it is not so clear but still the student took some lessons. Again, if a student says “Şu an olduğu gibi ara sıra yazma etkinlikleri yapıyorduk.” *(As we do now, we were doing occasional writing activities)*, we understand that they had some writing lessons, etc. Thus, it can be coded as (1).

Note: Again, sometimes it can be difficult to differentiate (1) and (2). If a student says “Çok az lisede aldım. Hiç yeterli değildi.” *(I took very few in high school. it was never*

enough), we coded it as (1). This is because, although the student says I took very few lessons, he still took some.

Question 2

Genel olarak İngilizce yazma performansınızı değerlendirir misiniz?
(Can you please evaluate your English writing performance in general?)

STEP 1: We created 5 categories to code the answers to this question.

The first category included answers which stated that they felt comfortable while writing/ were good at writing in English. For example, “Yaza yaza geliřtim. Farklı kelimeler ve cümleler kuruyorum.” (*I improved by writing and writing. I use different words and sentences.*) was coded as (1).

The second category included answers which stated that they need to deal with some issues to improve their writings. For example, “1 yıl öncesine kıyasen yazma performansım oldukça geliřti. Ama hâlâ istenilen kelime sayısına ulaşmakta zorlanıyorum.” (*Compared to 1 year ago, my writing performance has improved a lot. But I still have difficulty in reaching the desired word count.*) was coded as (2).

The third category included answers which stated that their writings need significant improvement. For example, “Fikirlerimi detaylı olarak yazmakta, açıklamakta, kelimeleri seçmekte zorlanıyorum.” (*I find it difficult to write down and explain my ideas in detail, and to choose words.*) was coded as (3).

The fourth category included “other” answers which cannot be classified because the answer was not clear. For example, “Değerlendiririm.” (*I will evaluate*) was coded as (4).

Additionally, there were some answers which mistakenly referred to the writing of the experiment rather than general English writing. These participants thought that the question asks them to evaluate the experimental writing task rather than their general writing experience. They were coded as irrelevant (5). For example, “Burada daha fikre dayalı yazılar yazdığımızdan çok seviyesi yüksek bir dil kullanmaktan ziyade basitçe fikrimi açıkladım.” (*Since we are writing more opinion-based articles here, I simply explained my opinion rather than using a very high-level language.*) was coded as (5).

The codes are also summarized in the following table (Table J2).

Table J2. The Categories and Samples for General English Writing Performance

Code	Category	Sample*
1	(Almost) No Difficulty/Good at Writing	<i>I improved by writing and writing. I use different words and sentences.</i>
2	A bit Difficult /Needs to Deal with some Issues	<i>Compared to 1 year ago, my writing performance has improved a lot. But I still have difficulty in reaching the desired word count.</i>
3	Difficult/Needs Significant Improvement	<i>I find it difficult to write down and explain my ideas in detail, and to choose words.</i>
4	Other	<i>I will evaluate</i>
5	Irrelevant (talks about the task at hand, etc.)	<i>Since we are writing more opinion-based articles here, I simply explained my opinion rather than using a very high-level language.</i>

*Since the Turkish of these samples were given above, I only added the translations to the table.

Note: Some participants gave a score on themselves. If they gave 7 or below out of 10, We coded them as (2), but if they gave 8 or above out of 10, We coded them as (1).

STEP 2: In the second step, we additionally coded the reason for having difficulty in writing by considering 12 different categories.

The codes are summarized in the following table (Table J3).

Table J3. The Categories and Samples for Subcategories of Experiencing Difficulty in General English Writing

Code	Category	Sample
0	No Difficulty /Irrelevant	Yetersiz. Çok tekrara düştüğümü düşünüyorum. Fikirlerim var ve açık olsa da destekleyemedim. (<i>Insufficient. I think I'm getting too repetitive. I had ideas but although they were open, I could not support them.</i>)
1	General English (may include vocabulary and grammar)	İçime pek sinmiyor. Dile hakim hissetmemek beni zorluyor. (<i>I don't like it much. It's hard for me not to feel confident while using the language.</i>)
2	Grammar	Grammerimin iyileşmesi lazım ama geri kalanlarda iyi olduğumu düşünüyorum. (<i>My grammar needs to improve, but I think I'm fine with the rest.</i>)
3	Vocabulary	İngilizce yazı yazmayı seviyorum. Cümle kurarken zorlanmıyorum ama kelime bilgim beni zorluyor. (<i>I love writing in English. I don't have difficulty in making sentences, but the vocabulary part is challenging.</i>)
4	General Writing Proficiency	İngilizce yazı yazmada tam olarak yeterli olduğumu düşünmüyorum çünkü kafamdakileri kağıda istediğim gibi dökemiyorum. (<i>I don't think I'm fully competent at writing in English because I can't get what's in my head down on paper as I want.</i>)
5	Content / Background Gap	Genellikle konuya göre değişiyor. Fikrimin çok olduğu bir konuda çok fazla şey yazabiliyorken fikrim yoksa tıkanıyorum. (<i>It usually depends on the subject. While I can write a lot about a subject that I have ideas for, if I have no idea, I get stuck.</i>)
6	Content Organization	İngilizce yazı yazmaktan hoşlanmıyorum ve iyi organize olmuş yazılar yazdığımı da düşünmüyorum. (<i>I do not like writing in English and I do not think that I write well-organized articles.</i>)
7	Planning	Hazırlık aşamasında değil de yazma kısmında daha çok detaylandırabiliyorum. Hazırlık kısmı stratejisini öğrenmem gerektiğini düşünüyorum. (<i>I can elaborate more in the writing, but not in the preparation phase. I think I need to learn the preparation part strategy.</i>)
8	Number of Words	İyi değil. Aynı kelimeleri tekrarladığımı ve kısa yazdığımı düşünüyorum. (<i>Not good. I think I repeat the same words and write short.</i>)
9	Duration	Zaman içine sıkıştırılmış her şeye kendimi gergin hissediyorum. Bu gerginlik doğru yazabilmemi de etkiliyor... (<i>I feel tense with everything compressed into time. This tension also affects my ability to write correctly...</i>)
10	Translation from Turkish	Bana İngilizce yazmak zor geliyor. Cümleleri Türkçe düşündüğüm için İngilizceye çevirmeye hâlâ tam alışamadım. (<i>I find it difficult to write in English. I still haven't quite gotten used to translating the sentences into English because I think in Turkish.</i>)
11	Other	Değerlendiririm. (<i>I evaluate.</i>)

Note: There were few statements which could fit into two different categories. In that case, we considered the category which had a larger scope. To illustrate, we coded “Bazen gramer ve vocabulary hataları yapabilirim ama her zaman anlatmak istediğimi en direkt biçimde anlatırım.” (*Sometimes I can make grammatical and vocabulary mistakes, but I always tell what I want to say in the most direct way.*) as (1) rather than (2) or (3), because it referred to some general English difficulty.

Note: If a statement was coded as 1 (no difficulty) or 5 (irrelevant) in the previous step, it was marked as 0 (No Difficulty/Irrelevant) because we looked for the underlying reason for the difficulty in this second step.

Note: It was difficult to differentiate code (4) and code (11) but considering the context of the question (general writing performance) was helpful. For example, if a participant said “ortalama derece” (*average rating*), it was concluded that the participant thinks that there are things related to the writing and needs to be dealt with and coded as (4).

Note: The statements like “Yetersiz. Çok tekrara düştüğümü düşünüyorum. Fikirlerim var ve açık olsa da destekleyemedim.” (*Insufficient. I think I'm getting too repetitive. I have ideas and even though they were open, I couldn't support them*) was coded as (0) because it refers to the experimental writing task rather than general writing.

Question 3

İlkokul ve lise dahil olmak üzere üniversite öncesinde Türkçe yazma dersleri aldınız mı? Aldıysanız ders içeriği, materyaller, ödevler, geribildirim vb. ile ilgili detay verebilir misiniz?

(Did you take Turkish writing courses before coming to the university, including primary and high school? If you did, can you give details about the course content, materials, assignments, feedback, etc.?)

Again, similar to its English version, we coded the answers to this question in three different categories. See the following table (Table J4) for the categories and samples please.

Table J4. The Categories and Samples for Taking Turkish Writing Lessons Before

Code	Category	Sample
1	Yes, I took writing lessons	Her aşamasında aldım. Daha çok kompozisyon şeklinde ödevlerdi. Öyle bir materyel hatırlamıyorum. / Ortaöğretimden itibaren aldım fakat yetersizdi. (<i>I took them at every stage. They were mostly assignments in the form of essays. I don't remember any such material. / I had it since secondary school, but it was insufficient.</i>)
2	No, I did not take writing lessons	Yazma dersleri almadım ama Türkçe ufak şiirler yazmayı denedim. (<i>I did not take writing lessons, but I tried to write small poems in Turkish.</i>)
3	Other (Not clear)	Özel olarak bir ders almadım (<i>I didn't take any separate/private lessons.</i>)

Note: Again, the most important differentiation may be between code (1) and code (3). If a student says something like “Birkaç defa kompozisyon adı altında ders ve ödevler verildiğini hatırlıyorum.” (*I remember a few times that lectures and assignments were given under the name of composition.*), we can code it as (1). We know that it is not so clear but still the student took some lessons. Again, if a student says “Lisede ve ortaokulda kompozisyon yazardık.” (*We used to write essays in the middle and high school*). It can be concluded that they had some writing lessons, etc. and it can be coded as (1).

Question 4

Genel olarak Türkçe yazma performansınızı değerlendirir misiniz?

(Can you please evaluate your Turkish writing performance in general?)

STEP 1: We considered the same five different categorizations (as in general English) for this question.

The first category included answers which stated that they felt comfortable while writing in Turkish. For example, “10 üzerinden 8 verebilirim Türkçe yazı performansına.” (*I can give 8 out of 10 for my Turkish writing performance.*) was coded as (1).

The second category included answers which stated that they need to deal with some issues to improve their writings. For example, “Aslında bayağıdır Türkçe kompozisyon yazmıyorum. Ama geliştirebilirim.” (*Actually, I haven't written compositions in Turkish*

for quite some time. But I can improve) and “İngilizceye göre daha iyi.” (*Better than English*) was coded as (2).

The third category included answers which stated that their writings need significant improvement. For example, “Türkçede paragraf yazarken zorlanıyorum. Bazen özellikle uzun paragraflarsa dikkatim dağılabiliyor.” (*I have a hard time writing paragraphs in Turkish. I can get distracted sometimes, especially with long paragraphs*) was coded as (3).

The fourth category included answers which cannot be classified because the answer was not clear. For example, “Uzun bir süredir Türkçe paragraf yazmadığım için tam performansımı bilmiyorum.” (*Since I haven't written a paragraph in Turkish for a long time, I don't know my exact performance.*) was coded as (4).

The last category included answers which cannot be classified because the answer was irrelevant. For example, there were some answers, such as e.g., “Hızlı ve aklıma gelenleri yazdım.” (*I wrote quickly and whatever came to my mind.*), which referred to the previous general English writing question. They were coded as (5).

They were also summarized in the following table (Table J5).

Table J5. The Categories and Samples for General Turkish Writing Performance

Code	Category	Sample
1	(Almost) No Difficulty/Good at Writing	<i>I can give 8 out of 10 for my Turkish writing performance.</i>
2	A bit Difficult /Needs to Deal with some Issues	<i>Actually, I haven't written compositions in Turkish for quite some time. But I can improve. / Better than English.</i>
3	Difficult/Needs Significant Improvement	<i>I have a hard time writing paragraphs in Turkish. I can get distracted sometimes, especially with long paragraphs.</i>
4	Other	<i>Since I haven't written a paragraph in Turkish for a long time, I don't know my exact performance.</i>
5	Irrelevant	<i>I wrote quickly and whatever came to my mind.</i>

*Since the Turkish of these samples were given above, I only added the translations to the table.

Note: As in the previous question, some participants gave a score on themselves. If they gave 7 or below out of 10, We coded them as (2), but if they gave 8 or above out of 10, We coded them as (1).

STEP 2: we further coded the reason for having difficulty in writing by considering 12 different categories (Table J6).

Table J6. The Categories and Samples for Subcategories of Experiencing Difficulty in General Turkish Writing

Code	Category	Some Samples
0	NoDifficulty /Irrelevant	Bu konuda iyi olduğumu söyleyebilirim. (<i>I can say that I am good at this.</i>)
1	General Turkish	**Türkçe'ye hakim olmadığım için zorlanıyorum. (<i>It's hard for me because I don't have a good command of Turkish.</i>)
2	Grammar	Türkçe yazma performansımı beğeniyorum ama dil bilgisine biraz daha dikkat edebilirim. (<i>I like my Turkish writing performance, but I can pay more attention to grammar.</i>)
3	Vocabulary	**Kelime bilgim yeterli değil. (<i>My vocabulary is not enough.</i>)
4	General Writing Proficiency	Aslında bayağıdır Türkçe kompozisyon yazmıyorum. Ama geliştirebilirim. (<i>Actually, I haven't written compositions in Turkish for quite some time. But I can improve.</i>)
5	Content /Background Gap	Bence iyiyim ama bazen ne yazacağımı bilmiyorum. (<i>I think I'm fine, but sometimes I don't know what to write.</i>)
6	Content Organization	(Türkçede de kompozisyonum zayıftı. Hiç anlamlı paragraflar oluşturamazdım.)* Önceki soruda da açıkladığım gibi pek iyi değil. ((<i>My composition was also poor in Turkish. I could not create meaningful paragraphs at all.</i>))* As I explained in the previous question, it is not very good.)
7	Planning	** Yazımı planlarken zorlanıyorum. (<i>I'm having a hard time planning my writing.</i>)
8	Number of Words	** Kelime sayısına ulaşmada zorlanıyorum. (<i>I'm having trouble reaching the word count.</i>)
9	Duration	Düşünerek yazdığımda uzun süreye sahipsem güzel yazarım. (<i>When I write by thinking carefully, if I have a long time, I write well.</i>)
10	Mechanics (E.g., Punctuation)	Türkçe ana dilim ama imlâ kurallarıyla pek bir alakam yok. Ayrıca bazen sesli harflerin noktalı kısımlarını unutabiliyorum. (<i>Turkish is my mother tongue, but I don't have much to do with spelling rules. Also, I can sometimes forget the dotted parts of the letters.</i>)
11	Other	Uzun bir süredir Türkçe paragraf yazmadığım için tam performansımı bilmiyorum. (<i>Since I haven't written a paragraph in Turkish for a long time, I don't know my exact performance.</i>)

*The answer in parenthesis was given to the previous question.

** My own samples.

Note: There were few statements which could fit into two different categories. In that case, we considered the category which has a larger scope.

Note: Again, some participants associated this question with the previous one and explicitly or implicitly referred to the answer they gave in that preceding question. This created some ambiguity in some of the answers. To solve this ambiguity and make an inference, we checked the answer to the previous question. For example, one of the participants said “Türkçede de kompozisyonum zayıftı. Hiç anlamlı paragraflar oluşturamazdım” (*My composition was also poor in Turkish. I could not create meaningful paragraphs at all*) as an answer to the previous (General English) question and he said “Önceki soruda da açıkladığım gibi pek iyi değil.” (*As I explained in the previous question, it is not very good.*) as an answer to this question. We concluded that this participant had some problems in content organization and coded accordingly... If it was not still clear, we coded the statement as “Not Clear/Other”.

Note: And again, if a statement was coded as 1 (no difficulty) or 5 (irrelevant) in the previous step, it was marked as 0 (No Difficulty/Irrelevant) because, in this second step, we looked for the underlying reason for the difficulty and did not consider those who do not experience difficulty, etc.

Question 5

Yazarken genelde Türkçe mi yoksa İngilizce mi düşünüyorsunuz? Detay verebilir misiniz?

(Do you usually think in Turkish or in English while writing? Can you give details?)

The answers to this question were coded into four categories. One of the important things in this question was the priority of a language. Sometimes the participants said that they thought both in English and Turkish but, for example, they generally started thinking in Turkish and this was followed by English. In those cases, they were coded as (1). The codes are summarized in the following table (Table J7).

Table J7. The Categories and Samples for Thinking in Turkish or in English while Writing in English

Code	Category	Sample
1	Thinking in Turkish	Türkçe düşünüyorum. Ondan sonra İngilizceye çevirmekle uğraşıyorum./ Önce Türkçe düşünüp sonra İngilizce düşünüyorum. (<i>I think in Turkish. Then, I try to translate it into English./ I think in Turkish first and then in English.</i>)
2	Thinking in English	Genelde İngilizce düşünüyorum çünkü Türkçe düşünürsem karıştırabileceğimi düşünüyorum. (<i>I usually think in English because if I think in Turkish, I think that I can mix it up.</i>)
3	Both English and Turkish	Karışık. Yazdığım konuya göre değişiklik gösteriyor. (<i>Mixed. It varies depending on the topic I'm writing.</i>)
4	Other	İngilizce yazarken Türkçe düşünürsem istediğim çeviriyi yapamıyorum. İngilizce düşündüğümde de kelime eksikliği hissediyorum. (<i>If I think in Turkish while writing in English, I cannot translate what I want. When I think in English, I feel a lack of words.</i>)

Question 6

Derslerde veya ders dışında ne tür yazma aktiviteleri yapıyorsunuz? Detay verebilir misiniz?

(What kind of writing activities do you do in or out of class? Can you give details?)

The answers to this question were coded into five categories. You can see the following table for the details (Table J8).

Table J8. The Categories and Samples for Writing Activities Done during or after the Lessons

Code	Category	Sample
0	None/Irrelevant	Şu an herhangi bir aktivite yapamıyorum./ Dizi ve film izlemeyi seviyorum. Arkadaşlarımla vakit geçiriyorum. <i>(I can't do any activities right now. / I like watching TV series and movies. I spend time with my friends.)</i>
1	Personal Writing	İzlediğim filmlerin değerlendirmelerini bir deftere, karakterlerin çizimiyle birlikte yorumluyorum. Günlüklerimi artık İngilizce tutmaya başladım. <i>(I interpret the comments of the movies I've watched in a notebook, along with the drawings of the characters. I started to keep my diaries in English now.)</i>
2	Self (Exam) Study	Derslerde konuyla ilgili şeyler yazıyorum. Ders dışında yazımı geliştirmek için kendime konu belirleyip onun üzerinde yazılar yazıyorum <i>(I write about the subjects in the lectures. In order to improve my writing outside of class, I set a topic for myself and write articles on it.)</i>
3	Related to School Work	Derslerde verilen yazı ödevlerini yapıyorum. Onun dışında yapmıyorum. <i>(I do the writing assignments given in the classes. Other than that, I don't.)</i>
4	Other	Writing aktiviteleri. <i>(Writing activities)</i>

Question 7

Müzik, tiyatro veya açık hava sporlarına özel bir ilgi duyuyor musunuz? Detay verebilir misiniz?

(Do you have a special interest in music, theater, or outdoor sports? Can you give details?)

Nine categories were created to code the answers to this question. They are summarized in the following table (Table J9).

Table J9. The Categories and Samples for Interest in Music, Theatre, and Outdoor Sports

Code	Category	Sample
0	No Interest	Eskiden duyuyordum. / Hiçbirine ilgi duymuyorum ama bu sporlar açısından düşünüp empati yapabiliyorum. (<i>I used to be interested in them. / I am not interested in any of them, but I can think and empathize in terms of these sports.</i>)
1	Music	Müziğe duyuyorum. Gitar çalıyorum ama diğer alanlara pek ilgim yok. (<i>I am interested in music. I play the guitar, but I'm not very interested in other fields.</i>)
2	Theatre	Tiyatroya biraz ilgim var. Onun dışında ilgi duymuyorum. (<i>I have some interest in theatre. Other than that, I'm not interested.</i>)
3	Outdoor	Açık hava sporlarına ilgi duyuyorum. Özellikle dağ tırmanışı gibi. (<i>I am interested in outdoor sports. Especially like mountain climbing.</i>)
12	Music & Theatre	Müzik ve tiyatro özel ilgi alanım. Müzik demek zaten hayat demek, tiyatro ise üniversite bölümüm. (<i>Music and theater are my special interests. Music means life, theater is my university department.</i>)
13	Music & Outdoor	Evet. Müzik ve açık hava sporlarıyla ilgilenebilirim. Piyano çalmak müzikte ilgilendiğim bir alan fakat spor dallarıyla daha çok ilgilenirim. (<i>Yes. I can be interested in music and outdoor sports. I am interested in playing piano, but I am more interested in sports.</i>)
23	Theatre & Outdoor	Tiyatroya karşı özel bir ilgim var. Lisede drama kulübüne katılmışım. Çocukluğumdan beri bir sürü spor yaptım. (<i>I have a special interest in theatre. I joined the drama club in high school. I've done a lot of sports since I was a kid.</i>)
123	Music & Theatre & Outdoor	Açık hava sporları dikkatimi çekse de bir tecrübem yok. Ancak geçmişte tiyatro kursuna gittim ve yakın geçmişten beridir piyano eğitimi alıyorum. (<i>Although outdoor sports caught my attention, I have no experience. However, I went to a theater course in the past and I have been taking piano lessons since the recent past.</i>)
4	Not Clear/Other	Evet ama daha önce yapma fırsatım olmadı. Bence iyi yönde zaman harcamak için çok iyi aktiviteler. (<i>Yes, but I haven't had a chance to do it before. I think they are very good activities to spend time in a good direction.</i>)

Note: Because they are highly related, the answers of the participants to the following, the 8th, question can create a context for this question, and vice versa. Thus, when there was an ambiguity, the coders checked the answers to both questions to finalize coding.

Note: after completing the coding process, I also merged the codes 1, 2, 3, 12, 13, 23, 123 to recode the data for a possible (interest versus non-interest) analysis. The new codes were as the following, 0: no interest, 1: interested in at least one of the clubs, and 2: not clear/other.

Question 8

Müzik, tiyatro veya açık hava sporu kuluplerine üye misiniz? Değilseniz olmayı düşünüyor musunuz? Detay verebilir misiniz?

(Are you a member of music, theater, or outdoor sports clubs? If you are not, do you plan to be? Can you give details?)

STEP 1: Again, the same categories were used to code the answers to this question. In the first step both active club membership or planning to become a club member (or taking lessons, etc.) were merged and coded together. One important thing the answers to this question revealed was that although the question was related to student club membership, many participants mentioned of their activities related to these three clubs. Because one of the aims of the question was to assess the background of the participants with respect to these activities, we considered these experiences and coded them accordingly. For example, if a participant said that he was in the orchestra, we coded him/her as an active and interested person in this area.

The codes are summarized in the following table (Table J10).

Table J10. The Categories and Samples for Being Active in Music, Theatre, and Outdoor Sports

Code	Category	Sample
0	Not Active/No Future Plan	Olmayı düşünmüyorum, zamanım yok. (<i>I don't plan to be, I don't have time.</i>)
1	Music	Müzik kulübüne üyeyim. (<i>I am a member of the music club.</i>)
2	Theatre	Okulumuzun tiyatro kulübüne üyeydim ama sağlık sorunlarından dolayı bıraktım. (<i>I was a member of our school's drama club but quit it due to health problems.</i>)
3	Outdoor	Spor kulübüne katılmak isterim. (<i>I would like to join the sports club.</i>)
12	Music & Theatre	Müzik ve tiyatro kulübündeydim. Ve bu alanlarda amatörce de olsa ilerlemek istiyorum. İki tutkulu alan (<i>I was in the music and drama club. And I want to advance in these areas, albeit amateurishly. two passionate fields</i>)
13	Music & Outdoor	(Açık hava sporlarını ve müziği çok severim)* Değilim ama üye olmak çok isterim. (<i>I love outdoor sports and music</i>)* <i>I'm not, but I would love to become a member.</i>)
23	Theatre & Outdoor	Tiyatro veya açık hava sporlarından birine üye olmayı düşünüyorum. (<i>I'm thinking of becoming a member of one of the theater or outdoor sports.</i>)
123	Music & Theatre & Outdoor	(Evet. Hepsine ilgim var.)* Hayır. Henüz üye değilim. Üye olmayı düşünüyorum. (<i>Yes. I'm interested in all of them.</i>)* <i>No. I am not a member yet. I am considering becoming a member.</i>)
4	Not Clear/Other	Değilim ancak düşünüyorum. (<i>I'm not, but I'm thinking.</i>)

*The answer in parenthesis was given to the previous question.

Note: Importantly, because this question was associated with the previous one, some participants explicitly or implicitly referred to the answer they had already given in the preceding question. This created some ambiguity in some of the answers. To solve this ambiguity and make an inference, we checked the answers to the previous question. To illustrate, one of the participants answered this question by saying “Değilim fakat uygun fırsat olursa kendimi geliştirmek için olmak istiyorum.” (*I am not, but I want to be in order to improve myself if there is a suitable opportunity*). Because this was an ambiguous statement, we needed to check the answer to the previous (club interest) question which was “Müziğe ilgi duyuyorum. Uzun zamandır enstrüman öğrenmek hayalim.” (*I am interested in music. It's been my dream to learn an instrument for a long time.*). The conclusion was that this participant was interested in music club and would

like to join this club if he/she can find an opportunity. Thus, it was coded as (1). Still, sometimes the context did not allow us to make a conclusion and in those cases, the answers were coded as not clear/other (4).

Note: after completing the coding process, I merged the codes 1, 2, 3, 12, 13, 23, 123 to recode the data for a possible (topic background) analysis. The new codes were as the following, 0: not an active/no future plan, 1: active or plans to be a member of at least one of the clubs, and 2: not clear/other.

STEP 2: In the second step, we focused and coded those who were/are active in any of these areas irrespective of whether they are (planning to be) the participants of a club or not, etc. Thus, those participants who are/were active in the area were isolated in this step. We used the same code numbers in the table above.

For example, “Bir dönem kendi ilçemin orkestrasındaydım.” or “Önceden enstrüman çalışıyordum.” (*I was in my own town's orchestra for a while. or I used to play an instrument.*) was coded as (1), “Özellikle tiyatroya ilgi duyuyorum. Bu alanda yakından ilgiliyim. Eğitim aldım ve sık sık seyreder, tiyatro metinleri okurum.” (*I am particularly interested in theatre. I am closely involved in this field. I was educated and I often watch and read theatrical texts.*) was coded as (2), and “Müzik ve tiyatro kulübündeydim. Ve bu alanlarda amatörce de olsa ilerlemek istiyorum. İki tutkulu alan” (*I was in the music and drama club. And I want to advance in these areas, albeit amateurishly. two passionate fields*) was coded as (12), but “Hayır ve düşünmüyorum.” (*No and I don't think*) was coded as (0).

Note: Similar to the previous part, we needed to check the answers to the 7th (club interest) question because many active experiences were reported in that question.

Note: Again, after completing the coding process, I merged the codes 1, 2, 3, 12, 13, 23, 123 to recode the data for a possible (active versus nonactive) analysis. The new codes were as the following, 0: no activities were done, 1: actively involved in at least one of the related activities, and 2: not clear/other.

APPENDIX K

POST WRITING (REFLECTIVE) QUESTIONS CODEGUIDE

Question 1

Yazarken okuyucularınızın (rektör ve kulüp üyeleri) bakış açısını dikkate aldınız mı? Aldıysanız, bu durum yazınız üzerinde etkili oldu mu? Nasıl? Detaylandırır mısınız?

(Did you consider the audience (rector and club members) perspective while writing? If you did, did it influence your writing? How? Can you give details?)

STEP 1: The answers to this question were first coded in six different categories. In this step, the focus was on the first part of the question: whether the reader perspective was taken or not.

To start with, the answers which included no PT such as “Hayır almadım çünkü eğer alsaydım heyecan yapardım ve yazım daha kötü olabilirdi.” (*No, I didn't because if I did, I would be thrilled, and my spelling could be worse.*) was coded as (0).

The ones which said “Yes” but did not specify whether they took the perspective of rector or student club members were coded as (1). For example, “Aldım ama çok da iyi bir şekilde görüş açısından bakamadım. Biraz zayıf kaldım.” (*I took it, but I could not consider the perspective very well. It was a little weak*) was coded as (1).

The next category included answers which focused specifically on the rector. For example, “Kulüp üyelerinden çok rektör olarak düşündüm çünkü belirli bir miktar var ve herkesin mutlu olması. Ayrıca okulun da zor durumda kalmaması gerekiyor. Bu şekilde düşünüp yazmaya başladım.” (*I thought of the rector rather than the club members because there is a certain amount [of money], and everyone is to be happy. In addition, the school should not be in a difficult situation. This is how I started thinking and writing*) was coded as (2).

On the other hand, the answers which focused on student club members, for example, “Kulüp üyelerinin bakış açısından aldım daha çok. Etkili oldu, onların bakış açılarını kullandım.” (*I took it more from the point of view of the club members. It was effective, I used their point of view*) was coded as (3).

There were also answers which focused on both the rector and the student club members. For example, “Her ikisinin de bakış açısını almaya çalıştım, öğrenciler alabilecekleri en iyi eğitimi ve ekipmanı istiyor; rektör de her kulübe elinden geldiği kadar destek olmaya çalışıyor.” (*I tried to take the perspective of both, students want the best education and equipment they can get; The rector tries to support every club as much as he can.*) was coded as (4).

The last category, other/not clear, included answers which could not fit any of the preceding categories. For example, “Daha çok öğrenci bakış açısıyla ele aldım.” (*I approached it more from a student perspective.*) was coded as (5)

Note: The main difference between the category (1) and (2), (3), (4) was that although the writer answered the question positively, he/she reported that there was not a clear reference to the rector or to the club members in category (1).

See the following table for the summary of the categories (Table K1).

Table K1. The Categories and Definitions for Taking Audience Perspective

Code	Category	Definition
0	No PT	No PT was reported.
1	Reader Perspective	The answer was positive but there was no explicit reference to the perspective of rector or student club members.
2	Rector Perspective	The answer was positive and there was an explicit reference to the rector, the rector’s perspective was taken.
3	Club Members Perspective	The answer was positive and there was an explicit reference to the club members, the club members’ perspective was taken
4	Both Rector and Club Members Perspective	The answer was positive and there was an explicit reference to the club members and the rector, their perspective was taken
5	Other/Not Clear	There seems to be a PT, but it is not clear.

Note: after completing the coding process, I merged the codes 1, 2, 3 and 4 to recode the data for a possible (PT versus no PT) analysis. The new codes were as the following, 0: no PT, 1: PT, and 2: other/irrelevant.

STEP 2: Then, in the second step, we focused on the second part of the question and coded the possible influence of PT on the writings.

We coded the answers with (0) if there was no influence of PT “on the text”, for example, “Yani fiyat açısından ele aldım diyebiliriz.” (*We can say that I took it in terms of price.*) or “Aldım ama çok da iyi bir şekilde görüş açısından bakamadım. Biraz zayıf kaldım.” (*I took it, but I could not consider the perspective very well. It was a little weak*). Here, the important thing is not whether the participant took the perspective of the reader(s) or not. This part of the question was coded in the previous step. The important thing is whether the participant mentioned of the influence of PT on the text. To illustrate, the second example includes PT but there is no implicit or explicit reference to the text.

We coded them with (1) if they had an influence on and contributed to “problem solution” (including being fair) as in “Aldım ama bütçe sınırlı olduğu için yapacak bir şey yoktu. Hepsine adil dağıtım yapmaya çalıştım ve öneriler sundum. Yani etkili oldu.” (*I took the perspective, but there was nothing to do because the budget was limited. I tried to give them a fair distribution and offered suggestions. So, it was effective*).

We coded them with (2) if they had an influence on and contributed to “the writing processes” (including the organization of content such as giving details) as in “Okuyucuların bakış açılarını dikkate aldım. Ve bu nedenle kararlarımı herkesin anlayabilmesi için yeterince örneklendirmeye çalıştım.” (*I considered the perspective of the readers. That's why I tried to illustrate my decisions enough for everyone to understand*).

We coded them with (3), other/not clear, if they had an influence but “does not fit into a clear category” as in “Kulüp üyelerinin bakış açısından aldım daha çok. Etkili oldu, onların bakış açılarını kullandım.” (*I took it more from the perspective of the club members. It was effective, I used their point of view*). Here, the participant says that PT had an influence, but he/she does not specify in which ways it was influential. It is different than (0), because although its influence is not clear, there is an influence of PT. In other words, the participant could take the perspective of the reader but does not refer to the text in (0). For example, “Aldım fakat müzik benim için her zaman birinci plandadır.” (*I took but music always comes first for me*) was coded as (0) because it suggests that the participant took the perspective of the readers but there is no indication of this PT on the text. Again, it is different than (1) in that there is a clear reference to problem solution (including being fair) in category (1).

Lastly, we coded them with (4) if there was an influence but the participant also reported some difficulties as in “Almaya çalıştım fakat bütçe onların istekleri karşısında çok yetersizdi. Çözüm bulmakta zorlandım.” (*I tried to take it, but the budget was very insufficient against their wishes. I had a hard time finding a solution*).

See the following table for the summary of the categories (Table K2).

Table K2. The Categories and Definitions for the Influence of Taking Audience Perspective

Code	Category	Definition
0	No influence of PT	The participant did not take the perspective of the reader at all, or he/she did not mention the influence of it on the text.
1	The Influence of PT on Problem Solution	The participant took the perspective of the reader and mentioned of its influence on problem solution (including being fair).
2	The Influence of PT on Text-Based Processes	The participant took the perspective of the reader and mentioned of its influence on writing processes (including organization).
3	The Influence of PT on Other Processes	The participant took the perspective of the reader and mentioned of its influence, but the answer was not detailed.
4	The Influence of PT with Accompanied Difficulty	The participant took the perspective of the reader and mentioned of its influence, but he/she also reported something negative.

Note: after completing the coding process, I merged the codes 1, 2, and 3 to recode the data for a possible (influence versus no influence) analysis. The new codes were as the following, 0: no PT influence, 1: PT influence, and 2: PT with accompanied difficulty.

Question 2

Yazma sorusunun anlaşılması kolay mıydı? Detay verebilir misiniz?

(Was the writing question easy to understand? Can you give details?)

We coded the answers to this question in three categories (Table K3).

The first category (1) included answers which stated that they had no difficulty to understand the question. For example, “Evet kolaydı. Rektör kendi başına karar vermek istemediği için herkese fikrini sormuş.” (*Yes, it was easy. Since the rector did not want to decide on his own, he asked everyone's opinion.*) and “Evet anlaşılabilirdi ve üzerine yazı yazılabilirdi çünkü tahmin edilebilir bir konuydu.” (*Yes, it was understandable and could be written on because it was predictable.*) were coded as (1).

The second category included answers which stated that it was a bit difficult to understand the question. For example, “Biraz zorlandım. Özellikle ilk kağıtta.” (*I struggled a little. Especially in the first paper.*) and “Orta dereceydi benim için. Kelime eksikliğinden dolayı takıldığım kısımlar oldu” (*It was mediocre for me. There were parts where I got stuck due to lack of words.*) were coded as (2).

The last category included answers which stated that it was difficult to understand the question. For example, “Zordu çünkü ilk defa gördüğüm İngilizce bir metin veya soru aklımı karıştırıyor.” (*It was difficult because when I see an English text or question for the first time, that confuses me.*) and “Zordu bence ya da ben dikkatimi veremediğim için tam anlayamadım” (*I think it was difficult or I couldn't fully understand because I couldn't pay attention.*) and “Zordu (kelimeler).” (*It was difficult (the words)*) were coded as (3).

Note: This question is about the prompt rather than the writing processes. Thus, if there is a sentence like “Evet kolaydı. Anladım ve yazmaya çalıştım. Sorun İngilizce yazarken konmu hakkında yeteri kadar yorum yapamamamdı.” (*Yes, it was easy. I understood and tried to write. The problem was that when I was typing in English, I couldn't comment sufficiently on the topic.*) It can be coded as (1) rather than (2). This is because although the participant says that he had some difficulties in writing, the writing question was easy to understand.

Table K3. The Categories and Samples for the Perceived Difficulty of the Prompt

Code	Category	Samples
1	(Almost) No Difficulty	<i>Yes, it was easy. Since the rector did not want to decide on his own, he asked everyone's opinion.</i>
2	Somewhat Difficult	<i>I struggled a little. Especially in the first paper.</i>
3	(Significant) Difficulty	<i>I think it was difficult or I couldn't fully understand because I couldn't pay attention.</i>

*Since the Turkish of these samples were given above, I only added the translations to the table.

Question 3

Metni yazarken zorlandınız mı? Detay verebilir misiniz?

(Did you experience difficulty while writing the text? Can you give details?)

STEP 1: Again, we coded the answers to this question in three categories (Table K4).

The first category included answers which stated that they had no difficulty to write the text. For example, “Hayır zorlanmadım. Düşünceler aklıma geldi. Sadece bazen bazı kelimelerin anlamlarını unuttum.” (*No, It was not so difficult. Thoughts came to my mind. It's just that sometimes I forget the meaning of some words.*) and “Hayır. Akıcı bir dille anlatmaya çalıştım.” (*No. I tried to explain in a fluent language.*) were coded as (1).

The second category included answers which stated that it was a bit difficult to write the text. For example, “Hayır zorlanmadım. Sadece yeterli seviyede İngilizcem olmadığı için düşünlerimi tam olarak yazıya dökemedim.” (*No, it was not difficult for me. I couldn't put my thoughts into writing just because I didn't have enough English.*) and “İngilizce yazmak biraz zorladı ama konu açısından bir metin yazmak kolaydı.” (*Writing in English was a bit of a challenge, but writing a text was easy in terms of the subject.*) were coded as (2).

The last category included answers which stated that it was difficult to write the text. For example, “Tüm kulüpleri seçmek istediğimden bütçe ayırmaya çalışırken zorlandım.” (*I had a hard time trying to allocate a budget as I wanted to choose all the clubs.*) and “Evet zorlandım çünkü ele alınması istenen fikre göre 250 kelime çok fazlaydı bence.” (*Yes, I had a hard time because I think 250 words was too much for the idea to be addressed.*) were coded as (3).

Note: There can be some exceptions but one of the key terms which can be assigned with code (2) is “biraz” (*somewhat*). There are some answers such as “Hayır, çok zorlanmadım. Sadece kelime sınırını doldurmak biraz zorlayıcı oldu.” (*No, it wasn't very hard. Just filling the word limit was a bit of a challenge.*). Here, the participant says that he did not have lots of difficulties, but it was a bit difficult to write a long enough passage. This must be coded as (2) rather than (1). Again, the following sentence “Biraz zorlandım. Karar veremedim. Havada kalmış olabilir.” (*I struggled a little. I could not decide. It might be up in the air.*) can be coded as (2) rather than (3).

Table K4. The Categories and Samples for the Perceived Difficulty in the Main Writing

Code	Category	Samples
1	(Almost) No Difficulty	<i>No. I tried to explain in a fluent language.</i>
2	Somewhat Difficult	<i>Writing in English was a bit of a challenge, but writing a text was easy in terms of the subject.</i>
3	(Significant) Difficulty	<i>Yes, I had a hard time because I think 250 words was too much for the idea to be addressed.</i>

*Since the Turkish of these samples were given above, I only added the translations to the table.

STEP 2: We further coded the reasons for having difficulty in 12 different categories.

Please see the following table for the categories and samples (Table K5).

Table K5. The Categories and Samples for the Sources/Subcategories of the Perceived Difficulty in the Main Writing

Code	Category	Sample
0	No Difficulty	Çok zorlanmadım. Düşüncelerimi paylaşmak benim için çok zor olmadı. (<i>It wasn't very hard. It wasn't very hard for me to share my thoughts.</i>)
1	General English	Metni yazarken bazı zamanlar zorlandım çünkü konuyla ilgili fikirlerimi İngilizce anlatmakta yer yer yetersiz kaldım. (<i>While writing the text, I had some difficulties at times because my English was sometimes insufficient in expressing my ideas on the subject.</i>)
2	Grammar	İngilizce grameri kullanırken biraz zorlandım. (<i>I had some difficulty using English grammar.</i>)
3	Vocabulary	Kelime stoğum az olduğu için biraz zorlandım. (<i>I had a little difficulty because my vocabulary level was low.</i>)
4	General Writing Proficiency	Biraz, yazma konusunda kötü olduğum için o da. (<i>A little, that's because I'm bad at writing.</i>)
5	Content /Background Gap	Biraz zorlandım çünkü pek aşina olduğum konular değildi. (<i>I had a bit of a hard time because it was not a subject I was very familiar with.</i>)
6	Content Organization	Kafamda yazının planını yaparken biraz zorlandım. Metnim biraz karışık oldu. (<i>I had some difficulty in planning the text in my head. It was disorganized.</i>)
7	The Number of Words	Kelime sayısına ulaşmakta zorlandım. Ama olabildiğince her şeyi açıklamaya çalıştım. (<i>I had a hard time reaching the word count. But I tried to explain everything as much as possible.</i>)
8	Duration	Süre sıkıntısı olduğundan çok yaratıcı olamadım. (<i>I couldn't be very creative due to time constraints.</i>)
9	Taking Perspective	Herkesin bakış açısına göre yazmak kolay olmadı ama herkese göre yazdığımı düşünüyorum (<i>It was not easy to write from everyone's point of view, but I think I wrote for everyone.</i>)
10	Finding a Solution	Biraz zorlandım çünkü çözüm bulmak zordu. (<i>I struggled a bit because it was difficult to find a solution.</i>)
11	Other	Orta seviyede. Çok zorlanmadım. (<i>Average. It wasn't very hard for me.</i>)

To note: The most important difference between code (1) and code (4) is that the emphasis in code (1) is on general English. For example, “Metni yazarken bazı zamanlar zorlandım çünkü konuyla ilgili fikirlerimi İngilizce anlatmakta yer yer yetersiz kaldım.” (*While writing the text, I had some difficulties at times because my English was sometimes insufficient in expressing my ideas on the subject.*) was coded as (1) but “Metin yazarken İngilizce yazdığımı için biraz zorlandım.” (*While writing the text, I*

had a little difficulty because we were writing in English.) was coded as (4). If we saw phrases such as “ifade etmekte” (expressing) with no reference to “yazmak”, (to write) we coded it as (1).

Question 4

Metni yazdıktan sonra gözden geçirme şansı bulabildiniz mi? Detay verebilir misiniz?
(Did you have a chance to review the text after writing it? Can you give details?)

To code the answers to this question, we created four categories. They were “Yes, I did” (1), “No, I did not prefer” (2), “No, there was a constraint” (3), “Other/Not Clear” (4). Please see the following table for the categories and samples (Table K6).

Table K6. The Categories and Samples for Reviewing after Writing

Code	Category	Definition	Sample
1	Yes, I did	The participant found time and revised the text.	Erken bitirdiğim için gözden geçirme şansım oldu. (Since I finished early, I had a chance to review it.)
2	No, I did not prefer	The participant had time but did not prefer revising the text.	Gerek duymadım. (I didn't need it.)
3	No, there was a constraint	The participant could not revise because of (e.g., time) constraint.	Hayır, o kadar vakit bulamadım. (No, I didn't have that much time.)
4	Other / Not Clear	The answer does not fit any of the preceding categories or it is not clear.	Süre fazla geldi kelime sayısına göre. (The time was longer than needed for the number of words.)

Question 5

Sizce yazdığınız bu metin nasıl değerlendirilecek? Detay verebilir misiniz?
(How do you think this text you wrote will be evaluated? Can you give details?)

To code the answers to this question, after revising the statements, four categories were created. They were the evaluation of “PT” (1), the evaluation of the “Ideas” of

participants (2), the evaluation of the passages by an “English Instructor (Teacher)” (3), and “Other/Not Clear” (4). Please see the following table for the categories and samples (Table K7).

Table K7. The Categories and Samples for the Predictions Made about the Evaluation of Main Writing

Code	Category	Sample
1	PT	Öğrencilerin bakış açılarını değerlendiriyor olabilirler. <i>(They may be assessing students' perspectives.)</i>
2	Ideas	Anlamak belki zor olur ama hak verilecektir./ Yazdığım yazıyı kulüp üyeleri için yararlı bulabilirler. <i>(It may be difficult to understand, but it will be appreciated. / They may find my passage useful for club members.)</i>
3	English Writing (Teacher)	Bazı kelimeleri çok sık kullandım. Bu açıdan olumsuz değerlendirilebilir. <i>(I used some words too often. In this respect, it may be evaluated negatively.)</i>
4	Other/Not Clear	Tam olarak ne amaçla değerlendirileceğini anlayamıyorum. Psikolojik bir deney gibi geliyor. <i>(I don't understand exactly for what purpose it will be evaluated. Sounds like a psychological experiment.)</i>

Note: It may be important to report that the code (2) may include key terms which implicitly or explicitly refer to the “ideas”. For example, the following statement, “Fikirler toplanarak uygun bir şekilde değerlendirilecek.” *(Ideas will be collected and evaluated appropriately.)* and “Yazdığım yazıyı kulüp üyeleri için yararlı bulabilirler.” *(They may find my passage useful for club members.)* were to be coded as (2).

APPENDIX L

CODING STRATEGY USE DURING EF TASKS

One strategy use question for each EF (WM and AC) task was asked to collect the data from the participants. Please see the following questions, operationalization, and samples for the finalized coding guideline for strategy use during EF tasks.

Question 1

Okların yönünü (sol-sağ) seçtiğiniz dikkat testini yaparken herhangi bir strateji kullandınız mı? Kullandıysanız detay verebilir misiniz?

(Did you use any strategies as you completed the attention task during which you chose the direction of the arrows (left or right)? If you used, can you give details?)

The answers to this question were coded as “Task Fulfillment/No Strategy Use”, “Pattern Strategy Use”, “Focus on One Button Strategy Use”, “Other Strategy Use (if any)”, or “Other/Irrelevant” (Table L1). The coding was done as the following.

Task Fulfillment/No Strategy Use: If the participant mentioned the requirement of the task (e.g., "I focused on +" or "Stratejim doğru seçeneği bulmaktı" *(My strategy was to find the right option)* or "Diğer okları görmemeye çalıştım" *(I tried not to see the other arrows)*), it is coded as "0".

Pattern Strategy Use: If the participant thought that there was a pattern and tried to follow that to improve his performance as in "Genellikle okların belli bir sıralama halinde olduğunu düşünüp ona göre tahmin ederek kafamda planladım. Mesela önce sol sonra sağ ya da iki sol sonra sağ gibi", (I usually thought that the arrows were in a certain order and planned it in my mind by guessing it. For example, first left, then right, or two left and then right.) it is coded as “11”.

Focus on One Button Strategy Use: If the participant used “focusing only on a part” (right button or left button) strategy to improve his performance as in "Evet oklar henüz çıkmadan sağa veya sola kilitlenip bekledim" *(Yes, before the arrows came out, I locked to the right or left and waited.)*, it is coded as “12”. Here I understand that the participant focused on pressing either the right or the left buttons. But there are some answers which talk about visually focusing on the left or the right of the screen, they must be coded as irrelevant because it is about antisaccade task.

Other Strategy Use (if any): If the participant used a strategy to improve his performance but it cannot be classified as 11 or 12 as in “Ok yönünü sol sol sağ sağ diye hangi yönse sürekli içimden tekrar ettim.” *(I kept repeating the direction of the arrow in my mind, whichever direction it was ... left, left, right, right ...)*, it is coded as “13”.

Other/Irrelevant: If the answer is not clear or relevant, for example, if the participant was confused and talked about antisaccade task (e.g., "Tek bir yöne baktım" (*I looked in one direction*)) or no conclusion can be made from the answer (e.g., "Evet" (*Yes*)) it is coded as "2". These examples were coded as Other/Irrelevant because the looking activity in the first example is related to antisaccade rather than flanker task and there is an ambiguity in the latter (it can be part of the task, etc.).

Table L1. The Categories and Samples for the Reported Strategy Use in Flanker Task

FLANKER TASK		
Code	Label	Samples
0	Task Fulfillment/No Strategy Use	<i>My strategy was to find the right option.</i>
11	Pattern Strategy Use	<i>I usually thought that the arrows were in a certain order and planned it in my mind by guessing it. For example, first left, then right, or two left and then right.</i>
12	Focus on One Button Strategy Use	<i>Yes, before the arrows came out, I locked to the right or left and waited.</i>
13	Other Strategy Use (if any)	<i>I kept repeating the direction of the arrow in my mind, whichever direction it was ... left, left, right, right.</i>
2	Other/Irrelevant	<i>I looked in one direction.</i>

*Since the Turkish of these samples were given above, I only added the translations to the table.

Note: after completing the coding process, I merged the codes 11, 12 and 13 to recode the data for a possible (strategy versus no strategy) analysis. The new codes were as the following, 0: no strategy, 1: strategy use, and 2: other/irrelevant.

Question 2

Yıldız işareti çıktığında O veya Q harfini yakalayabilmek için gözünüzü sola veya sağa kaydırduğınız dikkat testini yaparken bir strateji kullandınız mı? Kullandıysanız detay verebilir misiniz?

Did you use any strategies as you completed the attention task while trying to catch the O or Q letters by moving your eyes to the left or right when an asterisk appeared on the screen? If you used, can you give details?

The answers to this question were coded as “Task Fulfillment/No Strategy Use”, “Pattern Strategy Use”, “Focus on One Side Strategy Use”, “Other Strategy Use”, and “Other/Irrelevant” (Table L2). The coding was done as the following.

Task Fulfillment/No Strategy Use: If the participant mentioned the requirement of the task (e.g., “Evet hep zıt yöne bakmaya çalıştım” (*Yes, I always tried to look in the opposite direction*) or “Stratejim harfin altında kısa çizgi olup olmadığını bulmaktı” (*My strategy was to find out if there was a mark under the letter*) or “kullandım, yıldız çıktığı anda diğer tarafa hızlıca baktım” (*I used it, as soon as the star came out, I took a quick look at the other side*)), it is coded as “0”.

Pattern Strategy Use: If the participant thought that there was a pattern and tried to follow that to improve his performance as in “yıldızın ne tarafa gelineceğini tahmin edip ters tarafına bakmaya çalıştım” (*I tried to guess which way the star would go and look at the opposite side.*), it is coded as “11”.

Focus on One Side Strategy Use: If the participant used a strategy to improve his performance and if this is not part of the task (e.g., “Gözümü hep kenarlara odakladım” (*I always focused my eyes on the edges*) or “Gözümü kaydırmadan, yine aynı şekilde tek bir yere odaklanarak yaptım.” (*I did it without moving my eyes, still focusing on one place in the same way.*)), it is coded as (12). If the participant talks about a pattern (e.g., “3 kere sağa 3 kere sola odaklandım.” (*I focused 3 times to the right and 3 times to the left*)) or reports focusing on only one side, it can be coded as “12” but, as usual, the context must be considered, as well.

Other Strategy Use (e.g., focusing only to the center of the screen): If the participant used a strategy to improve his performance but it can not be classified as 11 or 12 (e.g., “genel olarak sağ veya sola değil ekranın ortasına bakmayı tercih ettim, bu şekilde farketmem daha kolay oldu” (*I generally prefer to look at the center of the screen rather than to the right or left, this way it was easier for me to notice.*)), it is coded as “13”.

Other/Irrelevant: If the answer is not clear or relevant (e.g., “Gözümde kayma olduğu için onu kullandım.” (*I have an esophoria problem in my eye, I used it*)) it is coded as “2”.

Note: Sometimes there can be a reference to the previously answered question as in “Flanker testinin aynısını yaptım.” (*I did the same as in the flanker test*), for this and any

of the following codes these kinds of previous source of the references will also be considered for coding.

Table L2. The Categories and Samples for the Reported Strategy Use in Antisaccade Task

ANTISACCADE TASK		
Code	Label	Samples
0	Task Fulfillment/No Strategy Use	<i>My strategy was to find out if there was a mark under the letter.</i>
11	Pattern Strategy Use	<i>I tried to guess which way the star would go and look at the opposite side.</i>
12	Focus on One Side Strategy Use	<i>I always focused my eyes on the edges.</i>
13	Other Strategy Use (e.g., focusing only to the center of the screen)	<i>I generally prefer to look at the center of the screen rather than to the right or left, this way it was easier for me to notice.</i>
2	Other/Irrelevant	<i>I have an esophoria problem in my eye, I used it.</i>

*Since the Turkish of these samples were given above, I only added the translations to the table.

Note: after completing the coding process, I merged the codes 11, 12 and 13 to recode the data for a possible (strategy versus no strategy) analysis. The new codes were as the following, 0: no strategy, 1: strategy use, and 2: other/irrelevant.

Question 3

Karelerin yerlerini aklınızda tutarken gördüğünüz resimler simetrik mi değil mi diye karar verdiğiniz hafıza testini yaparken bir strateji kullandınız mı? Kullandıysanız detay verebilir misiniz?

Did you use any strategies as you completed the memory task during which you tried to evaluate whether the presented figures were symmetrical while simultaneously trying to keep the location of the squares in your mind? If you used, can you give details?

The answers to this question were coded as “Task Fulfillment/No Strategy Use”, “Coding/Grouping”, “Rehearsal”, “Visualization”, “Other Strategy Use (if any)”, “Other/Irrelevant” (Table L3). The coding was done as the following.

Task Fulfillment/No Strategy Use: If the participant mentioned something highly related to the symmetry assessment or location recall requirement of the task (e.g., “Üst üste katlanabilir mi diye düşünerek yaptım” (*I did it thinking whether it can be folded over*) or “Görsel hafızamı kullandım daha çok” (*I mostly used my visual memory*)), it is coded as “0”.

Coding/Grouping: If the participant used a coding/grouping the location of squares strategy (e.g., by assigning a number to them) to improve his performance (e.g., “karelerin yerlerini aklımda tutmak için aklımda numaralandırdım” (*To keep the places of the squares in my mind, I numbered them in my mind.*)), it is coded as “11”.

Rehearsal: If the participant used a rehearsal strategy (e.g., by repeating the location of the squares) to improve his performance (e.g., “aklımda sürekli tekrar etmeye çalıştım” (*I tried to repeat it over and over in my mind*)), it is coded as “12”.

Visualization: If the participant used a visualization strategy to improve his performance (e.g., “Simetriye değil karelerin yerine daha çok odaklandım ve kareleri bir resim çizer gibi kafamda tuttum.” (*I focused more on the squares rather than the symmetry and held the squares in my head like I was painting a picture.*)), it is coded as “13”. To note, there is a clear visualization here. Otherwise, if the participant had said “I used only my visual memory”, it would probably be coded as 0, because this is what the task asks.

Other Strategy Use (if any): If the participant used a strategy to improve his performance but it cannot be classified as 11, 12, or 13 (e.g., “Elimle karelerin yerlerini tuttum” (*I marked the place of the squares with my hand*)), it is coded as “14”.

Other/Irrelevant: If the answer is not clear or relevant, for example, if the participant was confused and talked about foster task (e.g., “Harflerden kelime oluşturmaya çalıştım” (*I tried to make words from letters*), or “Evet” (*Yes*)) it is coded as “2”.

Table L3. The Categories and Samples for the Reported Strategy Use in Symmetry Span Task

SYMMETRY SPAN TASK		
Code	Label	Samples
0	Task Fulfillment/No Strategy Use	<i>I mostly used my visual memory.</i>
11	Coding/Grouping	<i>To keep the places of the squares in my mind, I numbered them in my mind.</i>
12	Rehearsal	<i>I tried to repeat it over and over in my mind.</i>
13	Visualization	<i>I focused more on the squares rather than the symmetry and held the squares in my head like I was painting a picture.</i>
14	Other Strategy Use (if any)	<i>I marked the place of the squares with my hand.</i>
2	Other/Irrelevant	<i>I tried to make words from letters.</i>

*Since the Turkish of these samples were given above, I only added the translations to the table.

Note: after completing the coding process, I merged the codes 11, 12, 13 and 14 to recode the data for a possible (strategy versus no strategy) analysis. The new codes were as the following, 0: no strategy, 1: strategy use, and 2: other/irrelevant.

Question 4

Harfleri aklınızda tutarken matematik problemlerini çözdüğünüz hafıza testini yaparken bir strateji kullandınız mı? Kullandıysanız detay verebilir misiniz?

Did you use any strategies as you completed the memory task while trying to solve mathematical problems while simultaneously trying to keep the letters in your mind? If you used, can you give details?

The answers to this question were coded as “Task Fulfillment/No Strategy Use”, “Coding/Grouping”, “Rehearsal”, “Visualization”, “Other Strategy Use (if any)”, “Other/Irrelevant” (Table L1). The coding was done as the following.

Task Fulfillment/No Strategy Use: If the participant mentioned something highly related to the arithmetic problem or letter recall requirement of the task (e.g., "Temel matematik bilgileri zaten strateji kullanılmaya da bilir" (*Only the basic mathematical knowledge was required, no need to use a strategy.*) or “matematik işlemlerini yaparken çok düşünmemeye çalıştım zaten genel olarak basit işlemlerdi, daha çok harflere odaklandım" (*I tried not to think too much while doing math operations, they were generally simple operations, I focused more on letters.*)), it is coded as "0".

Coding/Grouping: If the participant used a coding or grouping the letters strategy (e.g., by creating a word) to improve his performance (e.g., "harflerden sonra sesli harf getirerek aklımda daha kalıcı bir şekle sokmaya çalıştım" (*I tried to put it in a more permanent form in my mind by adding a vowel after the letters.*)), it is coded as "11". There were few cases which includes both coding and rehearsal (e.g., "Harflerin hecelerini birleştirerek içimden tekrar ettim." (*I repeated it inside myself, combining the syllables of the letters.*)). In those cases, we decided on which one has the priority over the other, but if there was no specific condition coding came first.

Rehearsal: If the participant used a rehearsal strategy (e.g., by repeating the letters) to improve his performance (e.g. Matematik sorularını çözerken aynı zamanda harfleri sürekli içimden tekrar ettim." (*While solving the math problems, I repeated the letters in my mind all the time.*) or "Her adımda harfleri hafızamda baştan tekrar ettim" (*I replayed the letters in my memory in every step*)), it is coded as "12". To note, there were few cases which were obvious that the participant meant letters but wrote about the numbers (e.g., "İçimden sayıları tekrar ederek, (tekerleme gibi) matematik problemlerini çözdüm" (*I solved math problems by repeating the numbers inside my mind, like rhymes*)). Considering the context, I coded these as "2", as well.

Visualization: If the participant used a visualization strategy to improve his performance (e.g., "Harfleri aklımda tutarken kullandığım en etkili strateji çıkan harfin bana ilk çağrıştırdığı şeyi gözümde canlandırmak oldu" (*The most effective strategy I used while keeping the letters in my mind was to visualize the first thing the letter reminded me of.*)), it is coded as "13".

Other Strategy Use (if any): If the participant used a strategy to improve his performance but it cannot be classified as 11, 12, or 13 (e.g., "Harflere aldırmamaya çalıp matematik sorularına odaklandım" (*I tried to ignore the letters and focused on the math questions.*)), it is coded as "14".

Other/Irrelevant: If the answer is not clear or relevant, for example, if the participant was confused and talked about a strategy for the arithmetic part of the task (e.g., "Evet" (Yes)) it is coded as "2".

Table L4. The Categories and Samples for the Reported Strategy Use in Foster Operational Span Task

FOSTER OPERATIONAL SPAN TASK		
Code	Label	Samples
0	Task Fulfillment/No Strategy Use	<i>Only the basic mathematical knowledge was required, no need to use a strategy.</i>
11	Coding/Grouping	<i>I repeated it inside myself, combining the syllables of the letters.</i>
12	Rehearsal	<i>While solving the math problems, I repeated the letters in my mind all the time.</i>
13	Visualization	<i>The most effective strategy I used while keeping the letters in my mind was to visualize the first thing the letter reminded me of.</i>
14	Other Strategy Use (if any)	<i>I tried to ignore the letters and focused on the math questions.</i>
2	Other/Irrelevant	<i>Yes.</i>

*Since the Turkish of these samples were given above, I only added the translations to the table.

Note: after completing the coding process, I merged the codes 11, 12, 13 and 14 to recode the data for a possible (strategy versus no strategy) analysis. The new codes were as the following, 0: no strategy, 1: strategy use, and 2: other/irrelevant.

APPENDIX M

CODING/EVALUATING

SIMPLE BRAINSTORMING AND PERSPECTIVE TAKING NOTES

THE MAIN GUIDELINE

- 1) Perspective Taking (PT) is defined as “the active cognitive process of imagining the world from another’s vantage point or imagining oneself in another’s shoes to understand their visual viewpoint, thoughts, motivations, intentions, and/or emotions” (Ku et al., 2015, p. 94-95). In other words, “at the level of mental representation, the effect of active PT will be to create a merging of self and other” (Davis et al., 1996) and it can result in an increased overlap between the mental representations of the self and the other which can lead perspective takers to coordinate their behaviors (Galinsky et al., 2005).
- 2) Keeping the operationalization above in our mind, we need to look at the notes the participants taken during the 3 minutes PT task (after reading the vignettes but before writing the main essays) and evaluate them on the basis of the items from Davis (1980) scale.

The following items are adapted from the PT scale of Davis (1980) to code the written essays.

- (Item 1 & Item 3) Seeing or imagining things (e.g., needs and expectations) from other club members of the rectorate’s points of view.
 - (Item 6 & Item 4) Putting themselves into the shoes of the club members/rectorate and showing a tendency to listen to their arguments regardless of agreeing or not with them.
 - (Item 2 & Item 5) Recognizing that there are more than one part of the question/problem and looking at different sides of it.
 - (Item 7) Imagining how the writer would feel if he/she was a club member.
- 3) The main aim is to consider and code them if they obviously took the perspective of club members! We will give “1” if we see evidence of PT and “0” if we cannot. The target of the PT (club members) can be mentioned, but it is not necessary.
 - 4) To note, if we feel that there is a general people’s PT, we can still code them, but we need to mark them with the word “overall”. For example, “Bence okuldaki kulüpler öğrenciler için çok önemli ...” (*I think clubs are very important for*

students) (participant 233) would be marked as “overall”. Again, sometimes, the participant can show that he/she took the perspective of other club members rather than the target club members, we can again code this by giving “1” but need to mark with the word “overall”.

- 5) Some participants took notes only by writing down some key words (e.g., one or two words), in that case we can look for some words which can refer to personal experience of those club members like “having fun, friendship, feeling peaceful, confident, etc.” or some specific words which can refer to the perspective of those club members such as “campfire, night, sun, etc.”. If we feel that there is a perspective, we can categorize them as the cases perspectives are taken. However, we again need to mark them as exceptional cases “key terms” so that they can be separated during the analysis.
- 6) If evidence of PT is detected, we need to evaluate the level of PT on a seven points scale. If there is no evidence for PT, we will give “0”. If there is a PT, we can give a score between “1” and “7” (*1 = very little PT, 4 = an average amount of PT, or 7 = strong PT*).
- 7) Lastly, some researchers talked about different kinds of PT, imagine-other and imagine-self PT (Todd & Galinsky, 2014). For example, Todd, Bodenhausen, Richeson, and Galinsky (2011) asked the participants either to imagine the target’s perspective (imagine-other) or to imagine themselves as if they were in the same situation with the target (imagine-self). This study did not aim to manipulate this differentiation, and Todd and Galinsky (2014) reported that studies which directly compared imagine-self (imagining as if they were the other person) and imagine-other (imagining other’s mental states) PT generally found no differences, but we will take notes if we encounter the material which shows that it is taken from imagine-self perspective. For example, this sentence would be an example for “imagine self”: “we love music and want to have a piano so that we can improve our club members’ PT abilities”. On the other hand, this statement would be an example for “imagine-other” PT: “They love music and want to have a piano so that they can improve their club members’ PT abilities”. Thus, when the participants directly put themselves into the shoes of the club members, we can take notes as “imagine-self” to separate these from the “imagine-other” cases.

(To note, because the aim of this coding was to check the manipulation from a different perspective and they were not the main concerns of the study and both of them would show the evidence for PT, I would have explored the imagine-self and imagine other distinction only if the data had revealed enough samples (*at least 10-15*) from each category. However, the data did not reveal enough imagine-self cases, thus I merged all PT cases.)

APPENDIX N

CODING INTERACTIONAL MARKERS

As Hyland and Milton (1997) did, the metadiscourse engagement marker (EM) coding was done by two researchers⁹² who were familiar with metadiscourse literature and with the classification of metadiscourse markers.

First, the coders became sure that they understood the operationalization and categorization of EMs, defined as “the ways writers rhetorically acknowledge the presence of their readers in a text” (Hyland & Jiang, 2016; p. 30). This definition suggests that the writers can determine the level of presence of a reader by employing EM, such as note that, etc., and in this way, they can get their readers’ attention and make them part of the discourse. Hyland (2005) suggests that EMs can address the readers in five different ways. The first way is to use reader pronouns (*you*, inclusive *we*, *your*, etc.) (p. 151-152). Secondly, the writers can use “personal asides” which “briefly interrupt the argument to offer a comment on what has been said” (p. 152). Thirdly, they can ask questions to their readers. Another strategy they can use to engage their readers is employing directives (giving instructions, etc.). Lastly, as Hyland and Jiang (2016; p.31) reported, they can use the appeals to shared knowledge markers (e.g., obviously) to ask readers to accept what is commonly known, etc. To gain more insight into these five subcategories, each coder carefully read at least Hyland (2005, p.151-155) and Hyland and Jiang (2016), in which the details and examples about using EMs were given.

During coding, the second person “you” and “your” and the other pronouns (inclusive “we”, “us”, our”) were considered/coded separately, thinking that the second person pronouns (as the direct reader mentions) could be considered as the existence of strong audience awareness (Fang & Zhuang, 2022) and as the most explicit demonstration of the recognition of the audience (Hyland & Jiang, 2022). Additionally, considering the context of the experiment, because the inclusive EMs “we”, “our” and “us” might refer to only one of the parties of the audience (rectorate or clubs), they were also differentiated in coding. For example, we coded “we” as an inclusive EM. However, I categorized it as a marker which included the rectorate as in “we should choose outdoor sports club and theatre club. Because of the university budget.” *or in* “In our university, we always support theatre players, and your education is so important for us.” However, we coded it as an inclusive EM which included only student clubs in the statements as: “... Outdoor clup sports are seasonal. when we get the equipment we just use 2-3 month.”. Furthermore, there were cases that included both rectorate and clubs. Again, we coded them accordingly as markers that engage both the rectorate and the students. Lastly, we considered those inclusive markers which involved other students or people who were not the main audience of writings and coded

⁹² Special thanks to Gülşah Sobucalı for her valuable support as a second evaluator.

them separately (but will not include them in the main analysis). We did the same for the second-person pronouns “you” and “your”.

To note, one important thing to consider in EM coding is the co-text because the same item’s function may be different in different texts, etc. (Adel, 2006; Hyland, 2005). To illustrate, we had to consider the context rather than form, for example, to differentiate “inclusive we” from “exclusive we”: the pronoun “we” in “last year, we went to a holiday with my family” would not be coded as an EM because it had nothing to do with the engagement of the audience. The same procedure was to be followed for coding SM markers (exclusive “we” and “I”) (See Table N1).

Table N1. Self-Mention and EM Categories

Markers and their Subcategories		Target Words or Samples
SM		
	Typical Self-Mention	"I, my, me"
	Exclusive Self-Mention	Excluding "we, our, us"
EM		
Pronouns	Inclusive Rector	"We, our, us" including the rectorate
	Inclusive Clubs	"We, our, us" including the club member(s)
	Inclusive Rector and Clubs	"We, our, us" including the rectorate and club(s)
	Inclusive All	"We, our, us" including all parties (university students, etc.)
	Second Pronoun Rector	"You, your" including the rectorate
	Second Pronoun Clubs	"You, your" including the club(s)
	Second Pronoun Rector and Clubs	"You, your" including the rectorate and club(s)
	Second Pronoun All	"You, your" including all parties (students, etc.)
Directives	Direct Engagement	E.g., "we should consider buying ..." etc.
Asides	Personal Asides	E.g., "theatre club (not appreciated enough) must be ..."
Questions	Rhetorical Questions	E.g., "why do we want to invite a professional theatre player?"
Appeals	Appeals to shared knowledge	E.g., "of course, music is good for everyone ..."

One important issue to be considered in coding was the possible existence of more than one EMs in a single sentence. For example, “we” is one of the most frequently used EMs in this context, but sometimes it occurs together with other EMs, as in “That was cool question let me show my answer to you”. Some researchers, for example, Ho and Li (2018), coded them twice as EMs (e.g., one as a pronoun and one as a rhetorical question, as in this sentence “Would you be happy?” (p. 61)). However, I encountered some samples in Hyland (2005), such as: “Then, let us consider a reference field which has rigid ...” (p. 154) and “You should note in methods a and b...” (p.165).

To deal with this problem and to be consistent in coding, I planned to code EMs only once, even when they occurred more frequently in a single statement. For this purpose, I divided passages into T-Units (minimal terminable units), which are “defined as one main clause plus whatever subordinate clauses are attached to it or embedded within” (Hunt, 1966, p. 738) and which many researchers used in discourse analysis and metadiscourse analysis (e.g., Min, Paek, & Kang, 2019), such as audience moves in an argumentative essay (Hays, Brandt, & Chantry, 1988). For example, Cho et al. (2021) adopted Hunt’s definition and separated T-Units to code PT, and Intaraprawat and Steffensen (1995) adopted a very similar approach to analyze metadiscourse features. Following this approach, I divided the passages into T-Units⁹³ (Appendix O).

One last note is about the priority in considering the type of an EM within a T-Unit. When there was more than one EM, the priority was to code them in the following order. We were to look at rhetorical questions first, then directives, personal asides, appeals to shared knowledge, and lastly, reader pronouns. Regarding the reader pronouns, second-person pronouns (you and your) will be prioritized over inclusive first-person pronouns (we, us, and our). This is because second pronouns were considered the most explicit demonstration of the recognition of the audience (Hyland & Jiang, 2022) and signaled strong audience awareness (Fang & Zhuang, 2022). In addition, typical SM (I, me, my) had a priority over exclusive we.

Following the division of the text into T-units, I composed EM lists used in the field, particularly the ones created by Hyland (2005, p.222-223). The target SM markers listed were: “I we, me, my, our, mine, us, the author, the author's, the writer, the writer's”. The target EMs were: “(), ?, (the) reader's, add, allow, analyze, apply, arrange, assess, assume, by the way, calculate, choose, classify, compare, connect, consider, consult, contrast, define, demonstrate, determine, do not, develop, employ, ensure, estimate, evaluate, find, follow, go, have to, imagine, incidentally, increase, input, insert, integrate, key, let x = y, let us, let's, look at, mark, measure, mount, must, need to, note, notice, observe, one's, order, ought, our (inclusive), pay, picture, prepare, recall, recover, refer, regard, remember, remove, review, see, select, set, should, show, suppose, state, take (a look/as example), think about, think of, turn, us (inclusive), use, we (inclusive), you, your”. Some additional EMs listed in Hyland and Jiang (2016) were also considered. They were: “typical, typically, usual, routinely, apparently, as a rule,

⁹³ For the analysis, I divided the number of T-Units, which include the EMs, into T-Units, which do not have any EMs. And this was the EM score of a participant. Another alternative to T-Units could be to count each EM and normalize their occurrences per 1,000 words [(raw frequency/total number of words) x 1,000]. However, as these authors suggested, metadiscourse is reflected on a more clause-level or higher scope. Thus, the T-Unit approach was more appropriate.

common, commonly, conventional, conventionally, established, familiar, normally, obvious, obviously, of course, prevailing, prevalent, traditional, traditionally, ought to, has to, - -, the regular expression *it is adj. to V.*". Importantly, because metadiscourse categories are open to newly encountered items (Hyland, 2005, p. 27), this initial list was revised/expanded by including "undoubtedly and usually". Additionally, the markers "do not" and "pay" were not considered in the initial stage but considered while evaluating the passages. This was because they were used very frequently, and their usage almost hardly ever served as EMs. Additionally, I quite rarely encountered markers given in quotation marks, obviously used instead of another word, or used in an incomprehensible context. For example, it is not possible to understand whether the pronoun "your" in "because when they played about harder move they say I did Audience like to move that like me, your model." refer to the reader of the passage or part of the reported speech. I decided not to code these kinds of items as markers.

After establishing the fundamentals, the target markers were used for the initial detection of possible EM and SM markers. This served as a template for further coding. Using this template, I carefully checked every statement and marked the SM and EMs by considering the context and adopting Hyland (2005) and Adel's (2006) explicitness criteria⁹⁴. The second independent coder did the same, first, for around 10 percent, and after checking the consistency, for the whole data. Upon the completion of all coding, Cohen's κ was run to determine the level of agreement between the researcher and the second coder. The result revealed an acceptable level of agreement between the coders for both EMs ($\kappa = .773, p < .001$) and SM markers ($\kappa = .981, p < .001$). Finally, the coders came together, discussed, and resolved all the discrepancies.

⁹⁴ That is, only overt textual devices which could be clearly identified in the text were considered, and the exclamation marks, smiling face emojis, etc. were not considered.

APPENDIX O

CODING/SEPARATING T-UNITS

The Main Guideline

As is operationalized by Bardovi-Harlig (1992)⁹⁵, if independent clauses, with subjects and finite verbs, are connected by a linker (and, or, but, etc.), it can be counted as two T-Units (1). However, if other clause(s) are embedded within an independent clause, it can be counted as one T-Unit as in (2) or (3) or (4) or (5) or (6). Sometimes, there can be run-on sentences. In that case, they can be divided as in (7).

- (1) They want to buy an acoustic piano, and we have enough money. (2 T-units/1 sentence)
- (2) They want to buy an acoustic piano which is very expensive. (1 T-unit/1 sentence)
- (3) If we pay 70.000 for music club, then we should give 10.000 for 2 club and it's not fair. (1 T-unit/1 sentence)
- (4) However, there is a problem because we have only 80.000 liras. (1 T-unit/1 sentence)
- (5) The other alternative is that some club members give money for their clubs so that university feel more comfortable. (1 T-unit/1 sentence)
- (6) So, I think we should give 30.000 lira to this club. (1 T-unit/1 sentence)
- (7) "I like the movie we saw about Moby Dick, the white whale [1st T-Unit] the captain said if you can kill the white whale, Moby Dick, I will give this gold to the one who can do it [2nd T-Unit] and it is worth sixteen dollars [3rd T-Unit] they tried and tried [4th T-Unit] but while they were trying they killed a whale and used the oil for the lamps [5th T-Unit] they almost caught the white whale [6]. (LI English, 4th grade) (6 T units/1 sentence)" (Bardovi-Harlig (1992; p. 390).

One important case which affected my coding was related to the usage of "because" which can have a special status in native English, as well (Foster, Tonkyn, & Wigglesworth, 2000). In Turkish, it is completely legitimate to start a new sentence with "because" (*bu yüzden*) and to give the reason in this sentence separately. Because of this, native Turkish students who learn English tend to transfer this to their L2. There

⁹⁵ Bardovi-Harlig (1992) recognizes that there are some difficulties in calculating, for example, the coordination (loss of the data related to conjunction usage). However, this applies to syntactic analysis, and he offers a solution for the problem. This will not create any problem for coding EMs within T-Units in this study's context.

are many cases that the participants did this in their writings. Considering this, I coded the sentences which start with “because” as independent T-Units as in (8).

- (8) ... However, I think, 30.000 lira is enough for theatre club. Because, there are a lot of equipments in music and they can play this equipments and they can earn money. (4 T-unit/2 sentences)

Beyond this, I ignored the accuracy in punctuation and syntactic structure (e.g., subject-verb agreement). And if the writer intended to say something but clearly used an incorrect word/form as in (9) where “so that” used instead of “so”, or in (10) where “using” is used instead of “use”, I ignored the incorrect usage and coded them accordingly. Again, there are some sentences which lack “to be” verb, but the meaning is very obvious as in (11), I counted these as separate sentences, too.

- (9) But rectors have only 80.000 lira so that we must choose one club. (2 T-units/2 sentences)

- (10) More group members using this money for buy equipment and education. (1 T-unit/1 sentence)

- (11) In there, people do some organizing, social activities, seminars. It an active club for people. (2 T-units/2 sentences)

REFERENCES

- Abdollahzadeh, E., Amini Farsani, M., & Beikmohammadi, M. (2017). Argumentative writing behavior of graduate EFL learners. *Argumentation*, 31, 641-661.
- Adel, A. (2006). *Metadiscourse in L1 and L2 English*. Amsterdam: John Benjamins Publishing.
- Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods*, 17(4), 351-371.
- Alamargot, D., Caporossi, G., Chesnet, D., & Ros, C. (2011). What makes a skilled writer? Working memory and audience awareness during text composition. *Learning and Individual Differences*, 21(5), 505-516.
- Anderson, R. C., & Pichert, J. W. (1978). Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior*, 17(1), 1-12.
- Apperly, I. (2018). Mindreading and psycholinguistic approaches to perspective taking: Establishing common ground. *Topics in Cognitive Science*, 10(1), 133-139.
- Apperly, I. A., Simpson, A., Riggs, K. J., Samson, D., & Chiavarino, C. (2006). Is belief reasoning automatic? *Psychological Science*, 17(10), 841-844.
- Baddeley, A. D. (1986). *Working memory*. Oxford, England: Clarendon Press.
- Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417-423.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. Bower (Eds.). *Psychology of Learning and Motivation* (pp. 47-89). New York: Academic press.
- Bajpai, S., Bajpai, R., & Chaturvedi, H.K. (2015). Evaluation of inter-rater agreement and inter-rater reliability for observational data: an overview of concepts and methods. *Journal of the Indian Academy of Applied Psychology*, 41(3), 20-27.
- Bardovi-Harlig, K. (1992). A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly*, 26(2), 390-395.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, 21(1), 37-46.

- Beauvais, C., Olive, T., & Passerault, J. M. (2011). Why are some texts good and others not? Relationship between text quality and management of the writing processes. *Journal of Educational Psychology*, 103(2), 415.
- Becker, A. (2006). A review of writing model research based on cognitive processes. *Revision: History, Theory, and Practice*, 25-49.
- Belotto, M. J. (2018). Data analysis methods for qualitative research: Managing the challenges of coding, interrater reliability, and thematic analysis. *The Qualitative Report*, 23(11), 2622-2633.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Block, M. K., & Strachan, S. L. (2019). The impact of external audience on second graders' writing quality. *Reading Horizons: A Journal of Literacy and Language Arts*, 58(2), 68-93.
- Bolat, U. (2019). *Ergenlerin yaşadığı duygusal ilişkilerde öz duyarlılık, kişisel tepkisellik ve problem çözme becerilerinin incelenmesi* (Master Thesis, Necmettin Erbakan Üniversitesi).
- Bradbury-Jones, J. Taylor & O.R. Herber (2014). Vignette development and administration: a framework for protecting research participants. *International Journal of Social Research Methodology*, 17(4), 427-440.
- Breland, H., Lee, Y.-W., & Muraki, E. (2005). Comparability of TOEFL CBT essay prompts: Response-mode analyses. *Educational and Psychological Measurement*, 65(4), 577-595.
- Brown, J. S., McDonald, J. L., Brown, T. L., & Carr, T. H. (1988). Adapting to processing demands in discourse production: The case of handwriting. *Journal of Experimental Psychology: Human Perception and Performance*, 14(1), 45-59.
- Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic Bulletin & Review*, 16, 893-900.
- Bruen, J., & Kelly, N. (2017). Using a shared L1 to reduce cognitive overload and anxiety levels in the L2 classroom. *The Language Learning Journal*, 45(3), 368-381.
- Byrnes, H., Manchon, R. M. (2014). Task-based language learning: Insights from and or L2 writing: An introduction. In Byrnes, H. & Manchon, R. M. (Eds.), *Task-based language learning—Insights from and for L2 writing* (pp. 1-27). Amsterdam, The Netherlands: John Benjamins.
- Calvard, T. S. (2010). *Understanding perspective taking and its role in relation to teamworking and diversity* (Doctoral dissertation, University of Sheffield).
- Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods & Research*, 42(3), 294-320.

- Carvalho, J. B. (2002). Developing audience awareness in writing. *Journal of Research in Reading*, 25(3), 271–282.
- Chadwick, S., & Ralston, E. (2010). Perspective-taking in structured and unstructured online discussions. *International Journal of Teaching and Learning in Higher Education*, 22(1), 1-11.
- Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, 81(1), 1-8.
- Chase, B. J. (2011). *An analysis of the argumentative writing skills of academically underprepared college students* (Doctoral dissertation, Columbia University).
- Cheng, F. W., & Chen, Y. M. (2009). Taiwanese argumentation skills: Contrastive rhetoric perspective. *Taiwan International ESP Journal*, 1(1), 23-50.
- Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication*, 18(1), 80-98.
- Cho, Y., & Choi, I. (2018). Writing from sources: Does audience matter? *Assessing Writing*, 37, 25-38.
- Cho, M., Kim, Y. S. G., & Olson, C. B. (2021). Does perspective taking matter for writing? Perspective taking in source-based analytical writing of secondary students. *Reading and Writing*, 34(8), 2081-2101.
- Cho, M., Kim, Y. S. G., & Wang, J. (2022). Perspective taking and language features in secondary students' text-based analytical writing. *Scientific Studies of Reading*, 1-16.
- Cohen, M., & Riel, M. (1989). The effect of distant audiences on students' writing. *American Educational Research Journal*, 26(2), 143-159.
- Colman, D. E., Letzring, T. D., & Biesanz, J. C. (2017). Seeing and feeling your way to accurate personality judgments: The moderating role of perceiver empathic tendencies. *Social Psychological and Personality Science*, 8(7), 806-815.
- Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, 169, 323-338.
- Cowan, N. (2011). The focus of attention as observed in visual working memory tasks: Making sense of competing claims. *Neuropsychologia*, 46(6), 1401-1406.
- Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769-786.

- Crane, P. K., Narasimhalu, K., Gibbons, L. E., Pedraza, O., Mehta, K. M., Tang, Y., ... & Mungas, D. M. (2008). Composite scores for executive function items: demographic heterogeneity and relationships with quantitative magnetic resonance imaging. *Journal of the International Neuropsychological Society*, 14(5), 746-759.
- Crocetti, E., Van der Graaff, J., Moscatelli, S., Keijsers, L., Koot, H. M., Rubini, M., ... & Branje, S. (2016). A longitudinal study on the effects of parental monitoring on adolescent antisocial behaviors: The moderating role of adolescent empathy. *Frontiers in Psychology*, 7, 1726.
- Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency: The roles of cohesion and linguistic sophistication. *Journal of Research in Reading*, 35(2), 115-135.
- Cumming, A. (1989). Writing expertise and second-language proficiency. *Language Learning*, 39(1), 81-135.
- Davelaar, E. J., Hannon, E. M., & Richards, A. (2005). Strategy use in the operation span task. In *Proceedings of the Annual Meeting of the Cognitive Science Society: CogSci 2005* (pp. 2464). Stresa, Italy.
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, 10, 85.
- Davis, M. H., Conklin, L., Smith, A., & Luce, C. (1996). Effect of perspective taking on the cognitive representation of persons: a merging of self and other. *Journal of Personality and Social Psychology*, 70(4), 713.
- Davis, M. H., Soderlund, T., Cole, J., Gadol, E., Kute, M., Myers, M., & Weihing, J. (2004). Cognitions associated with attempts to empathize: How do we imagine the perspective of another? *Personality and Social Psychology Bulletin*, 30, 1625-1635.
- DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 124-151). New York: Cambridge University Press.
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135.
- Draheim, C., Tsukahara, J. S., Martin, J. D., Mashburn, C. A., & Engle, R. W. (2021). A toolbox approach to improving the measurement of AC. *Journal of Experimental Psychology: General*, 150(2), 242-275.
- Duan, C., & Hill, C. E. (1996). The current state of empathy research. *Journal of Counseling Psychology*, 43(3), 261.

- Dunlosky, J., & Kane, M. J. (2007). The contributions of strategy use to working memory span: A comparison of strategy assessment methods. *Quarterly Journal of Experimental Psychology*, 60(9), 1227-1245.
- Eemeren, F. H., Henkemans, A. F. S., & Grootendorst, R. (2002). *Argumentation: Analysis, evaluation, presentation*. London: Routledge.
- Engeler, A. (2005). *Psikopati ve antisosyal kişilik bozukluğu* (Doctoral dissertation, İstanbul University).
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, 11(1), 19-23.
- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Eds.), *The Psychology of learning and motivation* (pp. 145–199). New York, NY: Elsevier
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology: General*, 128(3), 309.
- Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education*, 88(6), 915-933.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a non-search task. *Perception & Psychophysics*, 16(1), 143–149.
- Fang, Y., & Zhuang, G. (2022). Approaching the audience: engagement markers in Longinus' on the Sublime. *Language and Semiotic Studies*, 8(3), 86-105.
- Fayol, M. (2016). From language to text: The development and learning of translation. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research* (2nd ed.) (pp. 130–143). New York, NY: Guilford
- Ferretti, R. P., & Fan, Y. (2016). Argumentative writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed.) (pp. 301–315). New York: Guilford Press.
- Ferretti, R. P., & Lewis, W. E. (2019). Knowledge of persuasion and writing goals predict the quality of children's persuasive writing. *Reading and Writing*, 32, 1411-1430.
- Ferretti, R. P., Lewis, W. E., & Andrews-Weckerly, S. (2009). Do goals affect the structure of students' argumentative writing strategies? *Journal of Educational Psychology*, 101(3), 577.

- Ferretti, R. P., MacArthur, C. A., & Dowdy, N. S. (2000). The effects of an elaborated goal on the persuasive writing of students with learning disabilities and their normally achieving peers. *Journal of Educational Psychology*, 92(4), 694.
- Flanigan, J. R. (2021). *Evaluating the effects of a reader perspective-taking instructional strategy on the descriptive writing performance of children with a specific learning disability* (Doctoral dissertation, The University of Nebraska-Lincoln).
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365-387.
- Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition*, 43(2), 226-236.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21(3), 354-375.
- Fujii, Y. (2012). The effects of L1 on L2 writing and translation: A case study. *Journal of Modern Languages*, 22(1), 32-44.
- Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, 78(4), 708.
- Grabowski, J., Mathiebe, M., Hachmeister, S., & Becker-Mrotzek, M. (2018). Teaching perspective taking and coherence generation to improve cross-genre writing skills in secondary grades: A detailed explanation of an intervention. *Journal of Writing Research*, 10(2), 331-356.
- Gadermann, A. M., Schonert-Reichl, K. A., & Zumbo, B. D. (2010). Investigating validity evidence of the satisfaction with life scale adapted for children. *Social Indicators Research*, 96, 229-247.
- Galinsky, A. D., Ku, G., & Wang, C. S. (2005). Perspective-taking and self-other overlap: Fostering social bonds and facilitating social coordination. *Group Processes & Intergroup Relations*, 8(2), 109-124.
- Geng, H. (2017). *The effects of written feedback and revision-focus direction on student revision and writing improvement* (Doctoral dissertation, The University of Auckland).
- Gibbons, L. E., Carle, A. C., Mackin, R. S., Harvey, D., Mukherjee, S., Insel, P., ... & Alzheimer's Disease Neuroimaging Initiative. (2012). A composite score for executive functioning, validated in Alzheimer's Disease Neuroimaging Initiative (ADNI) participants with baseline mild cognitive impairment. *Brain Imaging and Behavior*, 6, 517-527.

- Gibson, W. (1950). Authors, speakers, readers, and mock readers. *College English*, 11(5), 265-269.
- Gilbert, K. (2004). A comparison of argument structures in L1 and L2 student writing. *Journal of Asian Pacific Communication*, 14(1), 55-75.
- Gilbert, S. J., & Burgess, P. W. (2008). Executive function. *Current Biology*, 18(3), 110-114.
- Gliner, J. A., Haber, E., & Weise, J. (1999). Use of controlled vignettes in evaluation: does type of response method make a difference? *Evaluation and Program Planning*, 22(3), 313-322.
- Godefroy, O., Gibbons, L., Diouf, M., Nyenhuis, D., Roussel, M., Black, S., ... & GREFEX Study Group. (2014). Validation of an integrated method for determining cognitive ability: implications for routine assessments and clinical trials. *Cortex*, 54, 51-62.
- Godinho, F., Maruta, C., Borbinha, C., & Pavao Martins, I. (2022). Effect of education on cognitive performance in patients with mild cognitive impairment. *Applied Neuropsychology: Adult*, 29(6), 1440-1449.
- Graham, K. M., & Eslami, Z. R. (2020). Does the simple view of writing explain L2 writing development?: A meta-analysis. *Reading Psychology*, 41(5), 485-511.
- Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, 99, 445-476.
- Hanson, R. K., & Scott, H. (1995). Assessing perspective-taking among sexual offenders, nonsexual criminals, and nonoffenders. *Sexual abuse: A Journal of Research and Treatment*, 7(4), 259-277.
- Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Randall (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1-27). Mahwah, NJ: Lawrence Erlbaum.
- Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, 29(3), 369-388.
- Hayes, J. R., & Berninger, V. (2014). Cognitive processes in writing: A framework. In Arfe, B., Dockrell, J., Berninger, V. (Eds.), *Writing development in children with hearing loss, dyslexia, or oral language problems: Implications for assessment and instruction* (pp. 3-15). New York: Oxford University Press.
- Hayes, J. R., & Chenoweth, N. A. (2007). Working memory in an editing task. *Written Communication*, 24(4), 283-294.

- Hayes, J.R., & Flower, L. (1980a). Identifying the organization of writing processes. In L.W. Gregg & E.R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–29). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hayes, J.R., & Flower, L. (1980b). The dynamics of composing: Making plans and juggling constraints. In L.W. Gregg & E.R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 31–50). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hays, J. N., Brandt, K. M., & Chantry, K. H. (1988). The impact of friendly and hostile audiences on the argumentative writing of high school and college students. *Research in the Teaching of English* 22(4), 391-416.
- Ho, V., & Li, C. (2018). The use of metadiscourse and persuasion: An analysis of first year university students' timed argumentative essays. *Journal of English for Academic Purposes*, 33, 53-68.
- Hoeffer, I. J., Van Knippenberg, D., Van Ginkel, W. P., & Barkema, H. G. (2012). Fostering team creativity: perspective taking as key to unlocking diversity's potential. *Journal of Applied Psychology*, 97(5), 982.
- Holliway, D. R. (2004). Through the eyes of my reader: A strategy for improving audience perspective in children's descriptive writing. *Journal of Research in Childhood Education*, 18(4), 334-349.
- Holliway, D. R., & McCutchen, D. (2004). Audience perspective in young writers' composing and revising. In L. Allal, L. Chanquoy, & P. Lamy (Eds.), *Revision of written language: Cognitive and instructional processes* (pp. 87-101). New York: Kluwer.
- Holt, L., Bobocel, D. R., & Chen, V. (2021). Delivering bad news fairly: Higher construal level promotes interactional justice enactment through perspective taking. *Journal of Organizational Behavior*, 42(6), 708-725.
- Hooper, N., Erdogan, A., Keen, G., Lawton, K., & McHugh, L. (2015). Perspective taking reduces the fundamental attribution error. *Journal of Contextual Behavioral Science*, 4(2), 69-72.
- Hoskyn, M., & Swanson, H. L. (2003). The relationship between working memory and writing in younger and older adults. *Reading and Writing*, 16(8), 759.
- Hruschka, D. J., Schwartz, D., St. John, D. C., Picone-Decaro, E., Jenkins, R. A., & Carey, J. W. (2004). Reliability in coding open-ended data: Lessons learned from HIV behavioral research. *Field Methods*, 16(3), 307-331.
- Hu, G., & Cao, F. (2015). Disciplinary and paradigmatic influences on interactional metadiscourse in research articles. *English for Specific Purposes*, 39, 12-25.
- Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes*, 20(3), 207-226.
- Hyland, K. (2005). *Metadiscourse*. London: Continuum.

- Hyland, K. (2010). Metadiscourse: Mapping interactions in academic writing. *Nordic Journal of English Studies*, 9(2), 125-143.
- Hyland, K., & Jiang, F. K. (2016). “We must conclude that...”: A diachronic study of academic engagement. *Journal of English for Academic Purposes*, 24, 29-42.
- Hyland, K., & Jiang, F. K. (2022). Metadiscourse choices in EAP: an intra-journal study of JEAP. *Journal of English for Academic Purposes*, 60, 101165.
- Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing*, 6(2), 183-205.
- Hunt, K. W. (1966). Recent measures in syntactic development. *Elementary English*, 43(7), 732-739.
- Intaraprawat, P., & Steffensen, M. S. (1995). The use of metadiscourse in good and poor ESL essays. *Journal of Second Language Writing*, 4(3), 253-272.
- Ito, F. (2004) The interrelationship among first language writing skills, second language writing skills, and second language proficiency of EFL university students. *JACET Bulletin* 39, 43-58.
- Ivanic, R. (1998). *Writing and identity: The Discoursal construction of identity in academic writing*. Amsterdam: John Benjamins.
- Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Rowley, MA: Newbury House.
- Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching*, 44(2), 137-166.
- Kadir Has Üniversitesi (n.d.). KHAS İngilizce Seviye Tespit ve Yeterlilik Sınavı (KHAS-STYS). Retrieved January 17, 2023, from <https://sfl.khas.edu.tr/khas-ingilizce-seviye-tespit-ve-yeterlilik-sinavi/>
- Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, 130(2), 169–183.
- Kellogg, R. T. (1996). A model of working memory. In C.M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57-72). Mahwah, NJ: Lawrence Erlbaum
- Kellogg, R. T. (1988). Attentional overload and writing performance: Effects of rough draft and outline strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(2), 355.

- Kellogg, R. T. (2001). Competition for working memory among writing processes. *The American Journal of Psychology*, 114(2), 175-191.
- Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research*, 1(1).
- Kellogg, R. T., Turner, C. E., Whiteford, A. P., & Mertens, A. (2016). The role of working memory in planning and generating written sentences. *Journal of Writing Research*, 7(3), 397-416.
- Kellogg, R. T., Whiteford, A. P., Turner, C. E., Cahill, M., & Mertens, A. (2013). Working memory in written composition: An evaluation of the 1996 model. *Journal of Writing Research*, 5(2), 159-190.
- Kemper, S., Kynette, D., Rash, S., O'Brien, K., & Sprott, R. (1989). Life-span changes to adults' language: Effects of memory and genre. *Applied Psycholinguistics*, 10(1), 49-66.
- Kim, Y. S. G. (2016). Direct and mediated effects of language and cognitive skills on comprehension or oral narrative texts (listening comprehension) for children. *Journal of Experimental Child Psychology*, 141, 101-120.
- Kim, Y. S. G. (2017). Why the simple view of reading is not simplistic: Unpacking the simple view of reading using a direct and indirect effect model of reading (DIER). *Scientific Studies of Reading*, 21(4), 310-333.
- Kim, Y. S. G. (2020). Structural relations of language and cognitive skills, and topic knowledge to written composition: A test of the direct and indirect effects model of writing. *The British Journal of Educational Psychology*, 90(4), 910-932.
- Kim, Y. S. G., & Graham, S. (2021). Expanding the Direct and Indirect Effects Model of Writing (DIEW): Reading–writing relations, and dynamic relations as a function of measurement/dimensions of written composition. *Journal of Educational Psychology*, 114(2), 215.
- Kim, K. J., & Pae, T. I. (2021). Examining the simultaneous effects of L1 writing, L2 reading, L2 proficiency, and affective factors on different task types of L2 writing. *Discourse Processes*, 58(7), 662-680.
- Kim, Y. S. G., & Park, S. (2019). Unpacking pathways using the Direct and Indirect Effects Model of Writing (DIEW) and the contributions of higher-order cognitive skills to writing. *Reading and Writing*, 32(5), 1319-1343.
- Kim, Y. S. G., & Schatschneider, C. (2017). Expanding the developmental models of writing: A direct and indirect effects model of developmental writing (DIEW). *Journal of Educational Psychology*, 109, 35-50.

- Kim, M., Tian, Y., & Crossley, S. A. (2021). Exploring the relationships among cognitive and linguistic resources, writing processes, and written products in second language writing. *Journal of Second Language Writing*, 53, 100824.
- Kim, Y. S. G., Wolters, A., Mercado, J., & Quinn, J. (2021). Crosslinguistic transfer of higher-order cognitive skills and their roles in writing for English-Spanish dual language learners. *Journal of Educational Psychology*, 114(1), 1-15.
- Knospe, Y. (2017). *Writing in a third language: a study of upper secondary students' texts, writing processes and metacognition* (Doctoral dissertation, Umea University).
- Kobayashi, H., & Rinnert, C. (1992). Effects of first language on second language writing: Translation versus direct composition. *Language Learning*, 42(2), 183-209.
- Kobayashi, H., & Rinnert, C. (2018). Composing processes. In J. I. Lontas (Eds.), *The TESOL encyclopedia of English language teaching* (pp. 1-9). United Kingdom: John Wiley & Sons, Inc.
- Kolovelonis, A., Pesce, C., & Goudas, M. (2022). The effects of a cognitively challenging physical activity intervention on school children's executive functions and motivational regulations. *International Journal of Environmental Research and Public Health*, 19(19), 12742.
- Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing*, 21(4), 390-403.
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks, CA: Sage.
- Ku, G., Wang, C. S., & Galinsky, A. D. (2015). The promise and perversity of perspective-taking in organizations. *Research in Organizational Behavior*, 35, 79-102.
- Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development*, 15, 309-328.
- Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science*, 22, 545-552.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.
- Lee, C., Ge, H., & Chung, E. (2021). What linguistic features distinguish and predict L2 writing quality? A study of examination scripts written by adolescent Chinese learners of English in Hong Kong. *System*, 97, 102461.

- Lee, C. J., & Goh, E. C. (2020). Using vignettes as a 'safe space' for low-income children to discuss sensitive topics in social work assessment. *Children and Youth Services Review*, 111, 104882.
- Leijten, M., Van Waes, L., Schriver, K., & Hayes, J. R. (2014). Writing in the workplace: Constructing documents using multiple digital sources. *Journal of Writing Research*, 5(3), 285-337.
- Li, J. (2006). The mediation of technology in ESL writing and its implications for writing assessment. *Assessing Writing*, 11(1), 5-21.
- Li, S. (2023). Working memory and second language writing: A systematic review. *Studies in Second Language Acquisition*, 1-33.
- Limpo, T., & Olive, T. (2021). Why should we be looking at the relationship between executive functions and writing? In Limpo, T., & Olive, T. (Eds.), *Executive functions and writing* (pp. 3-13). Oxford, UK: Oxford University Press.
- Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, 21(4), 861-883.
- MacArthur, C. A., Philippakos, Z. A., & Ianetta, M. (2015). Self-regulated strategy instruction in college developmental writing. *Journal of Educational Psychology*, 107(3), 855.
- Magnifico, A. M. (2010). Writing for whom? Cognition, motivation, and a writer's audience. *Educational Psychologist*, 45(3), 167-184.
- Majidi, A., Janssen, D., & de Graaff, R. (2021). The effects of in-class debates on argumentation skills in second language education. *System*, 101, 1-15.
- Martin, J. D., Tsukahara, J. S., Draheim, C., Shipstead, Z., Mashburn, C. A., Vogel, E. K., & Engle, R. W. (2021). The visual arrays task: Visual storage capacity or AC? *Journal of Experimental Psychology: General*. Advance online publication.
- Mason, L. H., & Shriner, J. G. (2008). Self-regulated strategy development instruction for writing an opinion essay: Effects for six students with emotional/behavior disorders. *Reading and Writing*, 21, 71-93.
- McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review*, 8, 299-325.
- McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist*, 35(1), 13-23.

- McCutchen, D. (2011). From novice to expert: Implications of language skills and writing-relevant knowledge for memory during the development of writing skill. *Journal of Writing Research*, 3(1), 51-68.
- McCutchen, D., Covill, A., Hoyne, S. H., & Mildes, K. (1994). Individual differences in writing: Implications of translating fluency. *Journal of Educational Psychology*, 86(2), 256.
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, 22(3), 276-282.
- McHugh, L., Barnes-Holmes, Y., & Barnes-Holmes, D. (2004). Perspective-taking as relational responding: A developmental profile. *The Psychological Record*, 54(1), 115-144.
- Midgette, E., Haria, P., & MacArthur, C. (2008). The effects of content and audience awareness goals for revision on the persuasive essays of fifth-and eighth-grade students. *Reading and Writing*, 21, 131-151.
- Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: An expanded sourcebook*. (3rd ed.). New York: SAGE Publications.
- Min, S., Paek, J. K., & Kang, Y. (2019). Exploring the Use of Hedges and Stance Devices in Relation to Korean EFL Learners' Argumentative Writing Qualities. *English Teaching*, 74(1), 3-23.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49-100.
- Mohr, P., Howells, K., Gerace, A., Day, A., & Wharton, M. (2007). The role of Perspective Taking in anger arousal. *Personality and Individual Differences*, 43(3), 507-517.
- Moore, N. S., & MacArthur, C. A. (2012). The effects of being a reader and of observing readers on fifth-grade students' argumentative writing and revising. *Reading and Writing*, 25, 1449-1478.
- Moreau, D., & Wiebels, K. (2021). Assessing change in intervention research: The benefits of composite outcomes. *Advances in Methods and Practices in Psychological Science*, 4(1), 1-14.
- Nawal, A. F. (2018). Cognitive load theory in the context of second language academic writing. *Higher Education Pedagogies*, 3(1), 385-402.

- Nussbaum, E. M., Kardash, C. M., & Graham, S. E. (2005). The effects of goal instructions and text on the generation of counterarguments during writing. *Journal of Educational Psychology*, 97(2), 157.
- Nussbaum, E. M., & Schraw, G. (2007). Promoting argument-counterargument integration in students' writing. *The Journal of Experimental Education*, 76(1), 59-92.
- O'Keefe, D. J. (1997). Standpoint explicitness and persuasive effect: A meta-analytic review of the effects of varying conclusion articulation in persuasive messages. *Argumentation and Advocacy*, 34(1), 1-12.
- Olive T. (2011). Working memory in writing. In Berninger V. W. (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology* (pp. 485-506). New York, NY: Psychology Press.
- Olive, T., Kellogg, R. T., & Piolat, A. (2008). Verbal, visual, and spatial working memory demands during text composition. *Applied Psycholinguistics*, 29(4), 669-687.
- Ong, J. (2013). Discovery of ideas in second language writing task environment. *System*, 41(3), 529-542.
- Ong, J. (2014). How do planning time and task conditions affect metacognitive processes of L2 writers? *Journal of Second Language Writing*, 23, 17-30.
- Ong, J., & Zhang, L. J. (2010). Effects of task complexity on the fluency and lexical complexity in EFL students' argumentative writing. *Journal of Second Language Writing*, 19(4), 218-233.
- Osborne, J. F., Henderson, J. B., MacPherson, A., Szu, E., Wild, A., & Yao, S. Y. (2016). The development and validation of a learning progression for argumentation in science. *Journal of Research in Science Teaching*, 53(6), 821-846.
- Phares, K. I. (2002). *The relationship between author and audience: Case study of a young-adult author and a student audience* (Doctoral Dissertation, Louisiana State University and Agricultural & Mechanical College).
- Rahimi, M. (2019). Effects of increasing the degree of reasoning and the number of elements on L2 argumentative writing. *Language Teaching Research*, 23(5), 633-654.
- Rahimi, M., & Zhang, L. J. (2018). Writing task complexity, students' motivational beliefs, anxiety and their writing production in English as a second language. *Reading and Writing*, 32(3), 761-786.

- Rahimi, M., & Zhang, L. J. (2021). Effects of an engaging process-genre approach on student engagement and writing achievements. *Reading & Writing Quarterly*, 1-17.
- Revesz, A. (2011). Task complexity, focus on L2 constructions, and individual differences: A classroom-based study. *The Modern Language Journal*, 95, 162-181.
- Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. Y. (2001). Influence of oral discussion on written argument. *Discourse Processes*, 32(2-3), 155-175.
- Reznitskaya, A., Kuo, L. J., Glina, M., & Anderson, R. C. (2009). Measuring argumentative reasoning: What's behind the numbers? *Learning and Individual Differences*, 19(2), 219-224.
- Rızaoğlu, F., & Gürel, A. (2022). Second language processing of English past tense morphology: The role of working memory. *International Review of Applied Linguistics in Language Teaching*, 60(3), 825-853.
- Richland, L. E., & Burchinal, M. R. (2013). Early executive function predicts reasoning development. *Psychological Science*, 24(1), 87-92.
- Roberts, A. L., Sumner, J. A., Koenen, K. C., Kubzansky, L. D., Grodstein, F., Rich-Edwards, J., & Weisskopf, M. G. (2022). Childhood abuse and cognitive function in a large cohort of middle-aged women. *Child Maltreatment*, 27(1), 100-113.
- Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics in Language Teaching*, 45(3).
- Robinson, P., & Gilabert, R. (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. *International Review of Applied Linguistics in Language Teaching*, 45(3), 161-176.
- Roen, D. H., & Willey, R. J. (1988). The effects of audience awareness on drafting and revising. *Research in the Teaching of English*, 22, 75-85.
- Roxßnagel, C. (2000). Cognitive load and perspective-taking: applying the automatic-controlled distinction to verbal communication. *European Journal of Social Psychology*, 30(3), 429-445.
- Ryskin, R. A., Benjamin, A. S., Tullis, J., & Brown-Schmidt, S. (2015). Perspective-taking in comprehension, production, and memory: An individual differences approach. *Journal of Experimental Psychology: General*, 144(5), 898-915.

- Qin, J., & Karabacak, E. (2010). The analysis of Toulmin elements in Chinese EFL university argumentative writing. *System*, 38(3), 444-456.
- Qin, J., & Liu, Y. (2021). The influence of reading texts on L2 reading-to-write argumentative writing. *Frontiers in Psychology*, 12, 655601.
- Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. In Rosenberg, S. (Eds.), *Advances in applied psycholinguistics. Volume 2. Reading, writing, and language learning* (pp. 142-175). Cambridge: Cambridge University Press.
- Schmidt, R.W. (1992). Psycholinguistic mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, 14, 357-385.
- Schneer, D. (2014). Rethinking the argumentative essay. *TESOL Journal*, 5(4), 619-653.
- Schneider, D., Lam, R., Bayliss, A. P., & Dux, P. E. (2012). Cognitive load disrupts implicit theory-of-mind processing. *Psychological Science*, 23(8), 842-847.
- Schoofs, D., Preuß, D., & Wolf, O. T. (2008). Psychosocial stress induces working memory impairments in an n-back paradigm. *Psychoneuroendocrinology*, 33(5), 643-653.
- Schoonen, R., Van Gelderen, A., Stoel, R. D., Hulstijn, J., & De Glopper, K. (2011). Modeling the development of L1 and EFL writing proficiency of secondary school students. *Language Learning*, 61(1), 31-79.
- Segalowitz, N. (2003). Automaticity and second languages. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 382-408). Malden, MA: Blackwell.
- Serafini, E. J., & Sanz, C. (2016). Evidence for the decreasing impact of cognitive ability on second language development as proficiency increases. *Studies in Second Language Acquisition*, 38(4), 607-646.
- Shah, P., & Miyake, A. (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.
- Silva, T. (1993). Toward an Understanding of the Distinct Nature of L2 Writing: The ESL Research and Its Implications. *TESOL Quarterly*, 27(4), 657-677.
- Skarlicki, D. P., & Turner, R. A. (2014). Unfairness begets unfairness: Victim derogation bias in employee ratings. *Organizational Behavior and Human Decision Processes*, 124(1), 34-46.
- Skehan, P. (1998). *A Cognitive approach to language learning*. Oxford: Oxford University Press.

- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510-532.
- Skilling, K., & Stylianides, G. J. (2020). Using vignettes in educational research: a framework for vignette construction. *International Journal of Research & Method in Education*, 43(5), 541-556.
- Smith, P. L. (1995). *Dimensions of audience in two composition classrooms: Reader/writer/text interactions* (Doctoral dissertation, Indiana University of Pennsylvania).
- Smith, I. H., Soderberg, A. T., Netchaeva, E., & Okhuysen, G. A. (2023). An examination of mind perception and moral reasoning in ethical decision-making: A mixed-methods approach. *Journal of Business Ethics*, 183(3), 671-690.
- So, C., & Joo, J. (2017). Does a persona improve creativity? *The Design Journal*, 20(4), 459-475.
- Stapleton, P., & Wu, Y. A. (2015). Assessing the quality of arguments in students' persuasive writing: A case study analyzing the relationship between surface structure and substance. *Journal of English for Academic Purposes*, 17, 12-23.
- Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research, and Evaluation*, 9(1), 4.
- Sunderman, G., & Kroll, J. F. (2009). When study-abroad experience fails to deliver: The internal resources threshold effect. *Applied Psycholinguistics*, 30(1), 79-99.
- Susser, B. (1994). Process approaches in ESL/EFL writing instruction. *Journal of Second Language Writing*, 3(1), 31-47.
- Syed, M., & Nelson, S. C. (2015). Guidelines for establishing reliability when coding narrative data. *Emerging Adulthood*, 3(6), 375-387.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (7th ed.). Boston, MA: Pearson.
- Tedick, D. J. (2002). Proficiency-oriented language instruction and assessment: Standards, philosophies, and considerations for assessment. In Minnesota Articulation Project, D. J. Tedick (Ed.), *Proficiency-oriented language instruction and assessment: A curriculum handbook for teachers (Rev Ed.)*. CARLA Working Paper Series. Minneapolis, MN: University of Minnesota, The Center for Advanced Research on Language Acquisition.
- Todd, A. R., Bodenhausen, G. V., Richeson, J. A., & Galinsky, A. D. (2011). Perspective taking combats automatic expressions of racial bias. *Journal of Personality and Social Psychology*, 100(6), 1027.

- Todd, A. R., & Galinsky, A. D. (2014). Perspective-taking as a strategy for improving intergroup relations: Evidence, mechanisms, and qualifications. *Social and Personality Psychology Compass*, 8(7), 374-387.
- Traxler, M. J., & Gernsbacher, M. A. (1993). Improving written communication through perspective-taking. *Language and Cognitive Processes*, 8(3), 311-334.
- Toulmin, S. E. (2003). *The uses of argument*. Cambridge University Press.
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, 37(3), 498-505.
- Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, 17, 635– 654.
- Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(6), 1302.
- Varghese, S. A., & Abraham, S. A. (1998). Undergraduates arguing a case. *Journal of Second Language Writing*, 7(3), 287-306.
- Vasylets, O., & Marín, J. (2021). The effects of working memory and L2 proficiency on L2 writing. *Journal of Second Language Writing*, 52, 100786.
- Vögelin, C., Jansen, T., Keller, S. D., Machts, N., & Möller, J. (2020). Organizational quality of ESL argumentative essays and its influence on pre-service teachers' judgments. *Cogent Education*, 7(1), 1760188.
- Wang, S. (2016). *An experiment of audience awareness effects on college students' argumentative writing* (Doctoral dissertation, The University of Nebraska-Lincoln).
- Wang, W., & Wen, Q. (2002). L1 use in the L2 composing process: An exploratory study of 16 Chinese EFL writers. *Journal of Second Language Writing*, 11(3), 225-246.
- Wardlow, L. (2013). Individual differences in speakers' Perspective Taking: The roles of executive control and working memory. *Psychonomic Bulletin & Review*, 20, 766-772.
- Watson, R., DeSanctis, G., & Scott Poole, M. (1988). Using a GDSS to facilitate group consensus: Some intended and unintended consequences. *MIS Quarterly*, 12, 463–478.

- Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.
- Weigle, S. C. (2005). Second language writing expertise. In K. Johnson (Eds.), *Expertise in second language teaching and learning* (pp. 128–149). Houndmills: Palgrave MacMillan.
- Weijen, D., Van den Bergh, H., Rijlaarsdam, G., & Sanders, T. (2009). L1 use during L2 writing: An empirical study of a complex phenomenon. *Journal of Second Language Writing*, 18(4), 235-250.
- Wen, Z., Lin, T. J., Glassman, M., Kim, S., Tilak, S., Nagpal, M., & Ha, S. Y. (2023). The development of early adolescents' social perspective taking through small-group discussion. *The Journal of Early Adolescence*, 0(0), 1-35.
- Westbrook, C., Baker, L., & Sowton, C. (2019). *Unlock 3 reading, writing and critical thinking* (2nd ed.). Cambridge: Cambridge University Press.
- Wilson, J., & While, A. E. (1998). Methodological issues surrounding the use of vignettes in qualitative research. *Journal of Interprofessional Care*, 12(1), 79-86.
- Wingate, U. (2012). 'Argument!' helping students understand what essay writing is about. *Journal of English for Academic Purposes*, 11(2), 145-154.
- Wolfe, C. R. (2011). Argumentation across the curriculum. *Written Communication*, 28(2), 193-219.
- Wolfe, E.W., & Manalo, J.R. (2004). Composition medium comparability in a direct writing assessment of non-native English speakers. *Language Learning & Technology*, 8(1), 53-65.
- Wollman-Bonilla, J. E. (2001). Can first-grade writers demonstrate audience awareness? *Reading Research Quarterly*, 36, 184–201.
- Woodbridge, A. (2017). *Cognitive perspective taking and audience awareness in second grade narrative writing* (Doctoral Dissertation, University of California).
- Xu, T. S., Zhang, L. J., & Gaffney, J. S. (2023). A multidimensional approach to assessing the effects of task complexity on L2 students' argumentative writing. *Assessing Writing*, 55, 100690.
- Yoon, S. O., Koh, S., & Brown-Schmidt, S. (2012). Influence of perspective and goals on reference production in conversation. *Psychonomic Bulletin & Review*, 19, 699-707.
- Zhang, Y. (2018). An investigation into the development of structure and evidence use in argumentative writing. *Theory and Practice in Language Studies*, 8(11), 1441-1448.

- Zhang, K. (2019). *Developing a rating scale for classroom assessment of the argumentative writing of Chinese EFL college students majoring in English* (Doctoral dissertation, University of Leicester).
- Zhu, Y., Mark Shum, S. K., Brian Tse, S. K., & Liu, J. J. (2016). Word-processor or pencil-and-paper? A comparison of students' writing in Chinese as a foreign language. *Computer Assisted Language Learning*, 29(3), 596-617.
- Zimmermann, R. (2000). L2 writing: Subprocesses, a model of formulating and empirical findings. *Learning and Instruction*, 10(1), 73-99.
- Zrimsek, C. (2013). *The writing of students with Asperger's syndrome* (Doctoral dissertation, University of New Hampshire).
- Zou, H. J., & Hyland, K. (2020). "Think about how fascinating this is": Engagement in academic blogs across disciplines. *Journal of English for Academic Purposes*, 43, 100809.