

INVESTIGATING TASKS IN THE EFL TEXTBOOKS IN THE TURKISH
ANATOLIAN HIGH SCHOOL CONTEXT

ÖZNUR DEMİRKAYA

A THESIS SUBMITTED

FOR

THE DEGREE OF MASTER OF ARTS

IN

TEACHING ENGLISH AS A FOREIGN LANGUAGE

İHSAN DOĞRAMACI BILKENT UNIVERSITY

ANKARA

SEPTEMBER 2022

Investigating Tasks in the EFL Textbooks in the Turkish Anatolian High School
Context

The Graduate School of Education

of

İhsan Doğramacı Bilkent University

by

Öznur Demirkaya

In Partial Fulfilment of the Requirements for the Degree of

Master of Arts

in

Teaching English as a Foreign Language

The Department of English Language Teaching
Ankara

September 2022

İHSAN DOĞRAMACI BILKENT UNIVERSITY

GRADUATE SCHOOL OF EDUCATION

Investigating Tasks in the EFL Textbooks in the Turkish Anatolian High School

Context

Öznur Demirkaya

September 2022

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Teaching English as a Foreign Language.

Asst. Prof. Dr. Necmi Akşit (Advisor)

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Teaching English as a Foreign Language.

Assoc. Prof. Dr. Perihan Savaş, Middle East Technical University (Examining Committee Member)

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Teaching English as a Foreign Language.

Assoc. Prof. Dr. Dilara Demirbulak, Yeditepe University (Examining Committee Member)

Approval of the Graduate School of Education

Prof. Dr. Orhan Arıkan (Director)

ABSTRACT**INVESTIGATING TASKS IN THE EFL TEXTBOOKS IN THE TURKISH
ANATOLIAN HIGH SCHOOL CONTEXT**

Öznur Demirkaya

MA in Teaching English as a Foreign Language

Advisor: Asst. Prof. Dr. Necmi Akşit

September 2022

This study was an attempt to investigate tasks in four English as a Foreign Language (EFL) textbooks designed by the Ministry of National Education (MoNE) and studied in the Anatolian high schools in Turkey in the 2020-2021 education year. Specifically, the study sought to identify task characteristics the activities possess, determine task types, and explore task complexity dimensions in the selected textbooks. The study also aimed to compare and contrast the grade levels with one another regarding these issues. In line with the research purposes, three checklists were generated, drawing from relevant frameworks for task characteristics (Ellis, 2011; Ellis & Shintani, 2014), task types (Ellis, 2003, 2011) and task complexity dimensions (Robinson, 2001a, 2001b, 2003, 2007, 2008, 2011, 2015; Skehan & Foster, 2001). Then, the activities in the textbooks were analysed through content analysis. The results demonstrated that only a small part of the activities in the textbooks qualify for a task, most of them failing to satisfy the “gap” and “outcome” criteria of taskness as defined by Ellis (2009, p. 223) and Ellis and Shintani (2014, p. 135). The findings related to the task types revealed that the sequencing of the task types across the grade levels, and therefore, proficiency levels, is not in line with the recommendations by Ellis (2020). Similarly, no systematic progression from simple to complex tasks as prescribed in the SSARC model by Robinson (2015) was detected within the textbooks. Implications of the study can be considered for practice and further research on task-based language teaching.

Keywords: task, task characteristics, task types, task complexity, task-based language teaching

ÖZET

TÜRKİYE’DE ANADOLU LİSESİ BAĞLAMINDA YABANCI DİL OLARAK OKUTULAN İNGİLİZCE DERS KİTAPLARINDAKİ GÖREVLERİN İNCELENMESİ

Öznur Demirkaya

Yabancı Dil Olarak İngilizce Öğretimi Yüksek Lisans Programı

Tez Danışmanı: Dr. Öğr. Üyesi Necmi Akşit

Eylül 2022

Bu çalışma, Milli Eğitim Bakanlığı (MEB) tarafından tasarlanan ve 2020-2021 eğitim öğretim yılında Türkiye'deki Anadolu liselerinde Yabancı Dil olarak okutulan dört İngilizce ders kitabında yer alan görevleri incelemeye yöneliktir. Çalışma, özellikle, seçilen ders kitaplarındaki etkinliklerde bulunan görev özelliklerini tespit etmeye, görev türlerini belirlemeye ve görev karmaşıklığı boyutlarını keşfetmeye çalışmıştır. Çalışma, aynı zamanda, sınıf düzeylerini bu konularda birbirleriyle karşılaştırmayı amaçlamıştır. Araştırma amaçları doğrultusunda, görev özellikleri (Ellis, 2011; Ellis & Shintani, 2014), görev türleri (Ellis, 2003, 2011) ve görev karmaşıklığı boyutları (Robinson, 2001a, 2001b, 2003, 2007, 2008, 2011, 2015; Skehan & Foster, 2001) için ilgili akademik çalışmalardan üç kontrol listesi oluşturulmuştur. Sonrasında, ders kitaplarındaki etkinlikler içerik analizi yoluyla analiz edilmiştir. Sonuçlar, ders kitaplarındaki etkinliklerin yalnızca küçük bir bölümünün görev olarak sayılabileceğini, çoğunun görevin, Ellis (2009, p. 223) ve Ellis ve Shintani (2014, p. 135) tarafından “boşluk” ve “sonuç” olarak adlandırılan kriterlerini karşılamadığını göstermiştir. Görev türlerine ilişkin bulgular, görev türlerinin sınıf düzeylerine ve dolayısıyla yeterlilik düzeylerine göre sıralanmasının, Ellis’in (2020) önerileriyle uyumlu olmadığını ortaya koymuştur. Benzer şekilde, ders kitaplarında, Robinson (2015) tarafından geliştirilmiş SSARC modelinde belirtildiği gibi basitten karmaşık görevlere doğru sistematik bir ilerleme tespit edilmemiştir. Çalışmanın sonuçları, görev temelli dil öğretimi üzerine uygulama ve araştırma için değerlendirilebilir.

Anahtar kelimeler: görev, görev özellikleri, görev türleri, görev karmaşıklığı, göreve dayalı dil öğretimi

ACKNOWLEDGEMENTS

I would like to express my feeling of gratitude to my advisor, Asst. Prof. Dr. Necmi Akşit for his continuous support and insightful comments to shape my thesis. It was always pleasure for me to prepare my thesis under his supervision. I also wish to express my thanks to the members of the committee, Assoc. Prof. Dr. Perihan Savaş and Assoc. Prof. Dr. Dilara Demirbulak for reading through my thesis out of their busy schedules and providing valuable suggestions.

I owe special thanks to the director of the School of Foreign Languages of Bartın University, Asst. Prof. Dr. Burcu Şentürk, for supporting me and giving me the permission to attend the Bilkent MA TEFL program.

I would like to thank Handenur Coşkun and Fatma Gümüşok who were involved in the coding process of this study. They were always kind enough to patiently answer my questions and assist me in coding. I am also grateful to Tuğçe Hamamcıoğlu and Elif Çinik for their valuable comments during the writing process.

My deepest thanks go to Betül Çimenli for showing unwavering support as a dear friend and a helpful colleague over the last three years. I am extremely thankful and indebted to her for offering me personal and professional support in each and every step of this thesis.

TABLE OF CONTENTS

ABSTRACT.....	iii
ÖZET.....	iv
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS.....	vi
LIST OF TABLES	xii
LIST OF FIGURES	xiv
LIST OF ABBREVIATIONS	xvi
CHAPTER 1: INTRODUCTION	1
Introduction.....	1
Background	3
Problem	7
Purpose.....	11
Research Questions	12
Significance.....	12
Definitions of Key Terms.....	15
CHAPTER 2: REVIEW OF RELATED LITERATURE	21
Introduction	21
Tasks and Task-based Language Teaching.....	21
Tasks and Task Characteristics	23
Activity.....	24
Task	24
Criterion 1	28
Criterion 2.....	29
Criterion 3.....	31

Criterion 4	31
Interpretation of Task Criteria by Erlam (2016) and Jaruteerapan (2020)	33
Non-task (Exercise)	39
Tasks versus Non-tasks	40
Task as a Unit of Syllabus	41
Task-based versus Task-supported Syllabus	44
Task Cycle	45
Task Types and Sequencing of Tasks	46
Focused versus Unfocused Tasks	48
Real-world versus Pedagogic Tasks	51
Input-based versus Output-based Tasks	52
Here-and-now versus There-and-then Tasks	55
Closed versus Open Tasks	55
Task Complexity and Sequencing of Tasks	58
Limited Attention Capacity	59
Task Difficulty Scheme	61
Code Complexity	62
Cognitive Complexity	62
Cognitive Familiarity	63
Cognitive Processing	63
Communicative Stress	63
Cognitive Hypothesis	65
Triadic Componential Framework	66
Task Complexity	68
Resource-directing Factors	69

Resource-dispersing Factors	74
Task Conditions	77
Task Difficulty	78
SSARC	79
Criticism against Triadic Componential Framework	81
Comparison of LAC and CH.....	88
Empirical Studies on Tasks.....	93
CHAPTER 3: METHOD	107
Introduction.....	107
Research Design.....	107
Context.....	109
English Language Education Policy in Turkey.....	109
Secondary Education Curriculum.....	110
EFL Textbooks Used in Secondary Education.....	114
Sampling	115
Instrumentation	116
Checklists.....	116
Checklist 1: Task Characteristics.....	117
Checklist 2: Task Types	120
Checklist 3: Task Complexity.....	121
Method of Data Collection and Analysis	128
Analysis of Language Activities for Task Characteristics	129
Criterion 1	132
Criterion 2.....	134
Criterion 3.....	136

Criterion 4	139
Analysis of Task Types	142
Focused versus Unfocused Tasks	143
Real-world versus Pedagogic Tasks	145
Input-based versus Output-based and Integrative Tasks	147
Here-and-now versus There-and-then Tasks	149
Closed versus Open Tasks	151
Analysis of Task Complexity Dimensions	152
Resource-directing Dimensions	155
There-and-then	155
Spatial Reasoning	156
Causal Reasoning	157
Intentional Reasoning	158
Perspective-taking	159
Resource-dispersing Dimensions	161
Single Task	161
Task Structure	161
Single Step	162
Prior Knowledge	163
SSARC Steps	165
Inter coder Reliability Analysis	166
Inter coder Reliability Analysis for Q1	167
Inter coder Reliability Analysis for Q2	170
Inter coder Reliability Analysis for Q3	172
CHAPTER 4: RESULTS	175

Introduction	175
Results of the Study	176
Tasks and Task Characteristics	176
Criterion 1	177
Criterion 2	177
Criterion 3	178
Criterion 4	179
Tasks and Non-tasks	181
Task Types and Sequencing of Tasks	182
Focused versus Unfocused Tasks	183
Real-world versus Pedagogic Tasks	183
Input-based versus Output-based and Integrative Tasks	184
Here-and-now versus There-and-then Tasks	185
Closed versus Open Tasks	186
Task Complexity and Sequencing of Tasks	189
Resource-directing Dimensions	190
Resource-dispersing Dimensions	191
SSARC Steps	192
CHAPTER 5: DISCUSSION	196
Introduction	196
Overview of the Study	196
Discussion of Major Findings	197
Tasks and Task Characteristics	197
Task Types and Sequencing of Tasks	204
Task Complexity and Sequencing of Tasks	211

Implications for Practice	216
Implications for Further Research.....	221
Limitations and Delimitations.....	225
REFERENCES.....	231
APPENDIX.....	255



LIST OF TABLES

Table		Page
1	Questions Relating to Coding of Criteria by Erlam	33
2	Coding Framework for Identifying Tasks by Jaruteerapan.....	35
3	Task Types Suggested for Different Proficiency Levels by Ellis	56
4	Task Difficulty Scheme by Skehan and Foster	62
5	Criteria for Sequencing Pedagogic Tasks by Robinson	67
6	Model English Language (9th – 12th Grades) Curriculum	113
7	The Organisation within Units in the High School EFL Textbooks	114
8	Frameworks Used for the Development of the Checklists in the Analysis Stages	116
9	The Framework for Coding Activities Based on the Study by Jaruteerapan (2020)	118
10	The Overview of the Analysis Stages Based on the Research Questions	128
11	The Frequencies and Percentages of the Activities Satisfying Criterion 1	177
12	The Frequencies and Percentages of the Activities Satisfying Criterion 2	178
13	The Frequencies and Percentages of the Activities Satisfying Criterion 3	178
14	The Frequencies and Percentages of the Activities Satisfying Criterion 4	179

15	The Frequencies and Percentages of the Tasks and Non-tasks in the EFL Textbooks	181
16	Focused and Unfocused Tasks Identified in the Textbooks	183
17	Real-world and Pedagogic Tasks Identified in the Textbooks...	184
18	Input-based, Output-based and Integrative Tasks Identified in the Textbooks	185
19	Here-and-now and There-and-then Tasks Identified in the Textbooks	186
20	Closed and Open Tasks Identified in the Textbooks	186
21	Resource-directing Dimensions Identified in the Output-based and Integrative Tasks.....	190
22	Resource-dispersing Dimensions Identified in the Output- based and Integrative Tasks.....	191
23	Task Complexity Steps Identified in the Output-based and Integrative Tasks Based on the SSARC Model	193

LIST OF FIGURES

Figure		Page
1	The Sequencing of Task Types as Proficiency Level Increases	57
2	A Sample Task Increased in Complexity Based on the Sequencing Principles of SSARC Model	80
3	Checklist 1 Used for Identifying Task Characteristics	117
4	Checklist 2 Used for Determining Task Types	120
5	Checklist 3 Used for Exploring Task Complexity	122
6	A Sample Unit Analysis for Task Characteristics	130
7	Theme 9 Activity 5b in <i>Count me in</i> for Grade 10	133
8	Theme 1 Activity 6 in <i>Teenwise</i> for Grade 9	133
9	Theme 2 Activity 2 in <i>Teenwise</i> for Grade 9	135
10	Theme 6 Activity 7 in <i>Silver Lining</i> for Grade 11	135
11	Theme 6 Activity 6a in <i>Teenwise</i> for Grade 9	137
12	Theme 7 Activities 7a-b in <i>Silver Lining</i> for Grade 11	138
13	Theme 10 Activity 5c in <i>Count me in</i> for Grade 10	140
14	Theme 9 Activity 5e in <i>Count me in</i> for Grade 12	141
15	Sample Analysis for the Identification of Task Types	142
16	Theme 2 Activity 3 in <i>Teenwise</i> for Grade 9	144
17	Theme 7 Activity 7 in <i>Silver Lining</i> for Grade 11	145
18	Theme 7 Activity 2b in <i>Count me in</i> for Grade 12	146
19	Theme 1 Activity 2 in <i>Teenwise</i> for Grade 9	146
20	Theme 10 Activity 1c in <i>Count me in</i> for Grade 10	148
21	Theme 8 Activity 6c in <i>Count me in</i> for Grade 12	148
22	Theme 1 Activity 5e in <i>Count me in</i> for Grade 10	149

23	Theme 6 Activity 6 in <i>Silver Lining</i> for Grade 11	150
24	Theme 9 Activity 2c in <i>Count me in</i> for Grade 12	150
25	Theme 5 Activity 6b in <i>Count me in</i> for Grade 10	151
26	Theme 8 Activity 6e in <i>Count me in</i> for Grade 10	152
27	Sample Analysis for the Exploration of Task Complexity Dimensions	153
28	Theme 3 Activity 9 in <i>Silver Lining</i> for Grade 11.....	156
29	Theme 5 Activity 8b in <i>Count me in</i> for Grade 10	157
30	Theme 8 Activity 3f in <i>Count me in</i> for Grade 10	158
31	Theme 4 Activity 2 in <i>Silver Lining</i> for Grade 11.....	159
32	Theme 3 Activity 7a in <i>Count me in</i> for Grade 10	160
33	Theme 10 Activity 5a in <i>Silver Lining</i> for Grade 11	161
34	Theme 7 Activities 7a-b in <i>Teenwise</i> for Grade 9	162
35	Theme 8 Activity 3-4 in <i>Teenwise</i> for Grade 9	163
36	Theme 1 Activity 7a-b in <i>Silver Lining</i> for Grade 11	164
37	Theme 3 Activity 7a in <i>Count me in</i> for Grade 10	165
38	The Percentages of the Activities Satisfying Task Criteria in the EFL Textbooks	180
39	The Percentages of the Tasks and Non-tasks in the EFL textbooks	182
40	The Percentages of the Task Types in the EFL Textbooks	188
41	The Distribution of Task Complexity Steps Based on the SSARC Model	194

LIST OF ABBREVIATIONS

C: Criterion

CAF: Complexity, Accuracy, Fluency

CCD: Cognitive/conceptual Demands

CH: Cognition Hypothesis

CEFR: Common European Framework of Reference

CL: Closed

CLT: Communicative Language Teaching

CoHE: Council of Higher Education

CON: Content

COS: Contextual Support

CR: Causal Reasoning

DIG: Discourse Genre

DYN: Dynamic

EBA: Educational Informatics Network

EFL: English as a Foreign Language

ELT: English Language Teaching

FO: Focused

HN: Here-and-now

IB: Input-based

INT: Integrative

IR: Intentional Reasoning

L1: First Language

L2: Second Language, Foreign Language

LAC: Limited Attentional Capacity

MoNE: Ministry of National Education

MUL: Multitasking

OB: Output-based

OP: Open

PE: Pedagogic

PK: Prior Knowledge

PPP: Presentation-practice-production

PT: Perspective-taking

Q: Question

RDir: Resource-directing

RDis: Resource-dispersing

RW: Real-world

SLA: Second Language Acquisition

SPSS: Statistical Package for the Social Sciences

SPA: Space

SR: Spatial Reasoning

SS: Single Step

SSARC: Stabilize, Simplify, Automatisé, Restructure, Complexify

ST: Single Task

STA: Static

STE: Steps

TBLT: Task-based Language Teaching

TCF: Triadic Componential Framework

TIM: Time

TOP: Topic

TS: Task Structure

TT: There-and-then

UN: Unfocused



CHAPTER 1: INTRODUCTION

Introduction

Task is among the fundamental concepts that have attracted attention in English language teaching (ELT) in recent years (Révész, 2019). It constitutes an essential component of syllabus design, language teaching, and assessment of individuals in both English as a foreign language (EFL) and English as a second language (ESL) contexts (Nunan, 2004). Having drawn considerable interest in the field of ELT since the introduction of the task-based approaches to language teaching in the 1980s, task is increasingly becoming an essential component of language textbooks.

Several definitions have been proposed to explicate the nature and features of tasks (Breen, 1989; Ellis, 2003; Long, 1985; Prabhu, 1987; Robinson, 2001a, 2001a; Skehan, 1998; Willis, 1996). Although there is no consensus on what a task as a concept denotes in the ELT literature, the definitions proposed for the concept of task generally refer to *task-as-workplan*, namely “the actual materials that constitute a task” (Ellis, 2018, p. 12), which is suggested to be the starting point that informs course design, testing, and research (Ellis et al., 2019). As specific design and implementation factors can predict the language use and activity emanating from task performance (Ellis, 2015), it could be argued that the design of a task-as-workplan to be implemented plays a crucial role in a language syllabus.

Among the critical issues in the task-based language teaching is the selection of tasks used in language textbooks to maximize the focus on form, implementation, and adaptation of tasks (Skehan, 1998). Another critical issue to be handled when designing a syllabus is the sequencing of tasks (Willis & Willis, 2007). As textbooks

are one of the primary sources for language teaching and learning that reflect curriculum and syllabus, they should follow a systematic method to optimize language learning when designing and sequencing tasks. Textbook designers need to determine what kinds of tasks to include in the textbooks as task demands are likely to affect the cognitive processes language learners are to get engaged in as well as their performance and learning process. Therefore, in addition to the design of tasks, the selection and sequencing of tasks are two critical procedures focused on in task-based syllabi (Ellis et al., 2019).

Examining task characteristics to determine task types is essential in making appropriate decisions on task design, selection, and sequencing. Task classification is valuable for offering variety, determining the kinds of tasks that serve the needs and preferences of different learner profiles, and providing language teachers with a framework to test and decide which types of tasks better support learners in learning a language (Ellis, 2003). Different task types may also provide distinct opportunities for developing communicative and linguistics competencies in L2.

Last but not least, another phenomenon investigated to make the right decisions on the design, selection and sequencing of tasks as workplan, namely the design materials, is the cognitive demands they pose, specifically *task complexity* (Robinson, 2001a, 2001b, 2003, 2005, 2007). It is assumed that manipulating task complexity dimensions to increase task complexity facilitates more opportunities for negotiation and interaction (Robinson, 2003, 2015), and increases attention to and memory for input during task interaction, thus promoting learning from L2 input and language development (Robinson, 2022). Hence, task complexity constitutes another

issue worthy of consideration to make appropriate decisions for optimally sequencing tasks in language textbook design.

Background

In the last decades, there has been an increasing focus on tasks and task-based approaches (Motteram & Thomas, 2010) in parallel with task-based language teaching (TBLT), which emerged as an alternative to Communicative Language Teaching (CLT) (Samuda & Bygate, 2008). The place of tasks in the syllabus has long been discussed (Ellis, 2003; Long, 1985; Robinson, 2001a, 2001b; Skehan, 1996; Willis, 1996), and tasks have been widely used in ELT textbooks (Rahimpour & Kiani, 2015).

In TBLT, it is recommended that task is conceptualized as a workplan (i.e., design materials) that informs the syllabus and course design (Ellis, 2018). On the other hand, there are the arguments of *Sociocultural Theory* and *Activity Theory* against the definition of task as a workplan since a task can be interpreted differently during task implementation process (Ellis, 2015) and may result in a somewhat different activity than expected (Ellis et al., 2019). Although the argument taken by these theories that task performances are constructed during the task implementation process is acknowledged, it is argued that it should still be recognized that particular tasks result in specific kinds of language behaviour as previous studies revealed that inherent features of tasks lead to “certain probabilistic process outcomes” (Ellis, 2018, p. 38). It is also maintained that there is a connection between the “task-as-workplan” and “task-as-process” (Ellis, 2009, p. 230), and they both can predict the kind of language task-takers produce to some extent (Ellis, 2012).

Although task is among the main concepts that have long been examined in the field of English Language Teaching (ELT), there has been no consensus on what

the concept of task involves with various operationalizations made in the literature (Ellis, 2003; Prabhu, 1987; Skehan, 1996; Willis, 1996). The definition of task put forward by Ellis (2009) and Ellis and Shintani (2014) constitutes one of the most cited frameworks for defining tasks, with its emphasis on *meaning*, a communicative *gap* to motivate language use, learners' use of their own *(non)linguistic resources* as well as the presence of a clearly defined *outcome*. As this conceptualization of task amalgamates some of the most highlighted characteristics of tasks, the task definition by Ellis (2009) and Ellis and Shintani (2014) is widely used to distinguish activities that qualify as tasks from non-tasks (i.e., exercises) in the ELT literature.

Task-as-workplan can be integrated into a syllabus in two different ways. As a “sub-component of the planning phase of curriculum development”, syllabus concerns the selection and grading of content (Nunan, 1988, p. 10). On the one hand, tasks can be the main component of a “task-based syllabus” in which they serve as the primary unit for selecting and sequencing course content (Lambert, 2019, p. 179). On the other, tasks may be the elements that do not essentially constitute the primary organising principle in “task-supported syllabus” (Révész, n.d.), where they basically function to support syllabus designed around other units of analysis such as “grammatical structures, lexis, topics, situations” (Lambert, 2019, p. 179). As syllabus is reflected in textbooks, one of the ways to determine whether a syllabus is task-based or task-supported is to identify tasks in the EFL textbooks, quantify them and calculate their proportion within these textbooks through content analysis. Such a query into a syllabus would help reveal whether a syllabus is identified in terms of tasks that determine and sequence instructional contents. Then it can be deduced whether this particular syllabus adopts a task-based approach or not (Ellis et al., 2019).

No matter if a syllabus is task-based or task-supported, the rationale behind task-based language teaching is not to engage the learners in performing tasks one after the other but to foster continuous learning and improvement (Willis, 1996). One way to ensure this is to determine task types (Ellis, 2011) for optimizing language learning through any syllabus design. Identifying and classifying the task characteristics guide textbook designers on how to sequence tasks in a syllabus (Robinson, 2007).

Several classifications have been developed for sequencing tasks based on specific features they gain depending on such factors as interactional criteria (Pica et al., 1993), operations (Willis, 1996), strategies (Nunan, 2004) and communicative and/or cognitive processes involved (Ellis, 2011, 2020; Ellis et al., 2019). Among these classifications, the task typology by Ellis (2011, 2020) and Ellis et al. (2019) has drawn considerable interest in the ELT literature with its dual categorization of tasks.

Another concern of task-based language teaching is the selection and sequencing of tasks based on task complexity. Adjusting task complexity by manipulating specific features of tasks has long been studied, drawing on cognitive, attention-driven approaches. Two prominent perspectives on task complexity *are Limited Attention Capacity (LAC) model (also known as Trade-off Hypothesis)* by Skehan (1996, 1998, 2009) and *Cognition Hypothesis (CH)* by Robinson (2001a, 2001b, 2003, 2005, 2007), both of which posit that the manipulation of task features affects the access to attentional resources of learners, their language performance, and thus, the development of interlanguage (IL), which denotes “the linguistic system of learner language produced by adults when they attempt meaningful communication using a language they are in the process of learning” (Tarone, 2013,

p. 2794). Even if Robinson and Skehan do not openly state that they solely focus on output-based skills (i.e., speaking and writing), the frameworks they propose have been taken as a basis in the studies which are mainly concerned with how the manipulation of task features is reflected in language production through the practice of these skills.

Drawing on the Levelt model targeting speaking performance in first language (L1), the *Limited Attention Capacity* (LAC) model consists of three categories in the task difficulty framework: “code complexity”, “cognitive complexity” and “communicative stress”, which relate to language, cognition and performance conditions, respectively (Skehan, 1996, p. 52; Skehan, 1998, p. 270; Skehan & Foster, 2001, pp. 194-195).

Robinson (2011, p. 6) proposes a *Triadic Componential Framework* (TCF) based on CH, which includes “task complexity”, “task conditions” and “task difficulty” dimensions concerning cognitive, interactive and learner factors, respectively. According to the CH, task complexity dimensions, which result from “the attentional, memory, reasoning, and other information processing demands imposed by the structure of the task on the language learner” (Robinson, 2001b, p. 29), must be taken as the sole determiner of syllabus design (Robinson, 2011). TCF offers a rationale for task sequencing based on the task complexity dimension so that learning is facilitated, and performance in the target language is satisfactorily actualized at different levels.

A close investigation into LAC and CH reveals some overlaps between these frameworks in terms of conceptualising some task demands affecting the cognitive complexity tasks have (Yang, 2014). Both frameworks are valuable for investigating task complexity; nevertheless, LAC by Skehan (1996, 1998) adopts a “moderate”

and “minimalist” approach to task complexity (Malicka, 2014a, p. 60) and does not provide a framework for task complexity as comprehensive as the TCF of CH.

Besides, Yang (2014) reports that previous researchers, having utilized Skehan’s framework for the task design phase, found it difficult to draw from. It should also be noted that LAC does not make suggestions specifically for syllabus design.

Different from LAC (Skehan, 1996, 1998; Skehan & Foster, 2001), TCF by Robinson (2001a, 2001b, 2003, 2005, 2007) examines *task complexity* in two dimensions: *resource-directing* and *resource-dispersing*, each of which consists of six sub-dimensions to explore the cognitive/conceptual and performative/procedural demands of tasks separately. Based on the demands these task complexity dimensions (*resource-directing* and *resource-dispersing*) pose, TCF proposes a model, “SSARC” (“stabilize”, “simplify”, “automatize”, “restructure”, “complexify”), to sequence tasks in a language syllabus (Robinson, 2015, p. 12). In this respect, TCF seems to provide a pretty comprehensive and detailed framework for task complexity (Malicka, 2014a) and suggestions for sequencing tasks. Appreciated also for being “the most recent and most robust” framework for task sequencing (Malicka, 2020, p. 644), TCF by Robinson, therefore, can be utilized as a primary framework for exploring task complexity in EFL textbooks.

Problem

Task, as one of the critical elements of language teaching, has long been investigated in the field of ELT; however, the concept of task is problematic as it still remains ambiguous to language teachers (East, 2018; Erlam, 2016; Jaruteerapan, 2021; Peng & Pyper, 2021), who need not only to incorporate tasks into their teaching practises but also to make implementational decisions regarding the types and complexity levels of tasks (Skehan, 1996.). Besides, it is argued that language

teachers have not yet been provided with suitable materials for task-based language teaching (Ellis, 2018).

Although the task-based approach is claimed to be adopted in the preparation of several ELT/EFL coursebooks, the previous query into these coursebooks revealed that they are not task-based but could qualify as task-supported at best (Ellis, 2018; Newton & Bui, 2017; Samuda & Bygate, 2008). It has been observed that the concept of task-based is not well-understood by coursebook designers as task-based is currently used to refer to almost any context, in which “so-called” tasks are used (Samuda & Bygate, 2008).

As far as the secondary school English curriculum in Turkey is considered, there is explicit emphasis is placed on task-based aspect of the curriculum, stating that how tasks are “presented in three different consecutive columns” for each grade level in the curriculum: *Functions and Useful Language, Language Skills and Learning Outcomes, and Suggested Materials and Tasks* (MoNE, 2018, p. 8), declaring that “[i]n all grades, communicative, experiential, and task-based language activities are promoted” (p. 14). Additionally, it includes a section called *Characteristics of English Language Teaching and Learning Environment (9th - 12th Graders)*, describing the attributes of materials/tasks, and how the assessment serves “...the language learning goals, materials, and tasks...” (p. 10). A number of textbooks have been developed with the aim of reflecting the curricular intentions; nevertheless, there are limited studies focusing the task-based nature of the curriculum (Ellis, 2009, 2012).

Surveying the studies conducted in Turkey in the last decade on task-based language teaching, the researcher observed that the studies were generally devoted to the efficiency of task-based language teaching, comparison of this teaching method

with other methods, investigation of activities in language materials from task-based method and proposal of tasks to be utilized in language teaching contexts.

Specifically, one group of studies compared task-based language teaching method with other teaching methods (Halıcı Page, 2016; Uzun & Ertok, 2020). Some studies investigated language teachers' (Mehmood, 2021; Kırtaş, 2016; Su-Bergil & Öcel, 2022) or students' (Akın, 2020; Uzun & Ertok, 2020) perceptions of task-based language teaching. Moreover, some studies explored its effect on anxiety (Kamalı, 2012), motivation (Asma, 2018; Halıcı Page, 2016; Kayır, 2018), self-efficacy beliefs (Akin, 2020; Değirmenci Mutlu, 2016), attitudes towards a course (Asma, 2018; Günel Şahan, 2019) as well as anxiety, motivation and self-confidence towards speaking English (Kamalı, 2012). As for the efficiency of task-based teaching, several studies were carried out to find out the effect of task based language teaching on writing (Ceylan, 2016), speaking (Ceylan, 2016; Şahin, 2019), vocabulary (Halıcı Page, 2016), vocabulary teaching process (İhtiyaryer, 2018), grammar knowledge (Yıldız, 2012), overall language skills (Akbulut, 2014) as well as academic success (Günel Şahan, 2019; Karakoç, 2016; Kayır, 2018).

It has been observed that most of the aforementioned studies were concerned with task implementation process. Only few described and analysed activities in language textbooks according to task-based language teaching method (Baklacı, 2018; Çelik, 2021; Rebahan, 2019; Yıldırım, 2021). A few studies proposed new language activities (Baklacı, 2018; Şimşek & Bakir, 2019) or lesson plans in line with task-based language teaching method (Selim, 2015). However, all of these studies concern the contexts in which Turkish is taught as L2.

To date, some research into the tasks in the EFL coursebooks was undertaken only by Günay (2007) in the primary school context in Turkey; however, no studies

conducted in the Turkish high school context were found when the Thesis Database of Council of Higher Education (CoHE) and DergiPark were explored. Therefore, the first gap in the ELT literature is due to the insufficient number of studies that explore to what extent the activities in the EFL coursebooks used in the Turkish high school context possess the task characteristics.

When the relevant literature is surveyed, it has been observed that little has been done to study the task types and their sequencing in the ELT coursebooks (Alemi et al., 2013; Alemi & Hesami, 2014; Günay, 2007; Huda et al., 2020; Yoon, 2017). Besides, previous studies did not use the task taxonomy proposed by Ellis (2003, 2011) as a point of reference for studying task types and their sequencing. The ones that used this taxonomy, on the other hand, generally examined one or two task types within dual categories to compare learner motivation, performance or language acquisition during task-as-process (Ahour & Shemshadsara, 2015; Alavinia et al., 2018; Hamavandy & Golshan, 2015; Lee, 2020; Rahimpour & Kiani, 2015; Shadabi et al., 2017). Therefore, a gap has been noticed that no study has been conducted to examine all the task types and their sequencing in the EFL coursebooks based on Ellis's (2003, 2011) framework.

Another problem lends itself to the insufficient number of studies on task sequencing based on task complexity (Motteram & Thomas, 2010). The importance of making principled decisions concerning the complexity dimensions of tasks before incorporating them into textbooks is emphasized (Samuda & Bygate, 2008; Vasylets et al., 2017). One way to accomplish this is to make decisions about task complexity by exploring the task features and what cognitive and performative demands they make. *Cognitive Hypothesis* (CH) by Robinson (2001a, 2001b, 2003, 2005, 2007) proposes a framework for exploring task complexity; however, it has been generally

used to study the task output with specific respect to the complexity, accuracy, and fluency (CAF) of the task-taker's L2 production or the interaction and negotiation that occur during task implementation process through experimental or semi-experimental research designs. To the best of the researcher's knowledge, only the studies by Barrot (2019) and Butler et al. (2018) directly attempted to explore the task complexity in EFL coursebooks based on the *Triadic Componential Framework* by Robinson.

In the Turkish context, only one study was set to examine task complexity in terms of language, cognitive involvement, and interactional demands in an EFL coursebook designed for primary school context (Günay, 2007). The other relevant studies in the Turkish context were conducted to delve specifically into the influence of manipulation of cognitive task complexity on written performance (Gökgöz, 2009; Sülükçü, 2010; Yıldız, 2017). Nevertheless, no study has been conducted to explore how task complexity is handled in the EFL textbooks for high school context based on Robinson's framework. Therefore, the lack of systematic analysis of the task complexity in the high school context in Turkey constitutes the last gap to be addressed in the present study.

Purpose

The overarching aim of this study is to provide a comprehensive analysis of the tasks in the English textbooks in the Anatolian high school context in Turkey. Specifically, the purpose of the present study is threefold.

First, it aims to determine to what extent the activities in the English textbooks used in the Anatolian High school context possess the task characteristics. Such a query is also expected to reveal the proportion of tasks in the textbooks. In

this way, it seeks to provide an insight into the role of tasks within the syllabi of the English textbooks in the Anatolian High school context in Turkey.

Second, the study aims to determine the task types used in the English textbooks in the Anatolian High school context, and how they are sequenced across the textbooks, drawing conclusions on how appropriate the textbooks are for the target learner groups in terms of the task types presented.

In addition, the present research is an attempt to explore how the output-based and integrative tasks (i.e., the tasks that engage learners in using any skills from both input-based and output-based tasks) are sequenced based on task complexity dimensions within and across the textbooks used in the Anatolian High school context. Specifically, this attempt intends to bring to the fore both cognitive/conceptual and performative/procedural demands the tasks in the textbooks make on language learners. Thus, it aims to discuss whether the textbooks follow a principled approach to sequencing tasks based on the demands in question.

Research Questions

The present study was initiated as an effort to answer the following research questions:

1. To what extent are the activities in the English textbooks used in the Anatolian High school context possess task characteristics?
2. How are task types sequenced in the textbooks?
3. How are task complexity dimensions of output-based and integrative tasks sequenced in the textbooks?

Significance

The present study hopes to contribute to the existing literature in the field of ELT in several respects. Overall, this study may extend the current literature by

providing a comprehensive analysis of tasks, task types and task complexity dimensions in the four EFL textbooks used in the Anatolian high schools in Turkey.

Specifically, the current study sheds light on to what extent the activities in the English textbooks used in the Anatolian high school context possess the task characteristics, and what the role of the tasks in the language textbooks are. The outcomes of this would provide the material developers with information about the current state of the tasks in the textbooks, and what alternative steps, if any, might be taken when developing textbooks. The study would also help language teachers gain insights into how to redesign activities to engage learners in meaningful tasks.

There could be different ways of diversifying and sequencing tasks, and this would lead to various consequences on learning the target language (L2). The present study offers and uses lenses to analyse tasks in the secondary school English textbooks in Anatolian High school context in Turkey. It places emphasis on how tasks could be sequenced across the textbooks prepared for different proficiency levels. One way to accomplish this is to determine the appropriate task types to incorporate into the textbooks targeting different proficiency levels. The findings of the study would reveal what types of tasks are presented and how they are sequenced in the EFL textbooks. They would inform material developers about the appropriacy of task types in the textbooks for each proficiency level. The data from the present study can also help language teachers make online decisions regarding the task types while teaching different learner groups.

Drawing on the predictions of the *Cognitive Hypothesis* (CH), the data from the present study may add to the previous research by revealing how the task complexity dimensions of the output-based and integrative tasks are sequenced in the EFL textbooks. It can also be helpful to make predictions for the task demands on

language learners during task implementation, which arise from cognitive/conceptual and performative/procedural factors. The data from the present study may provide suggestions for material developers to design and sequence tasks, which would engage learners in certain cognitive processes associated with specific task complexity dimensions. Language teachers would also be informed about the potential effects of manipulating tasks along different dimensions on learner performance and language development.

All things considered, the findings may offer valuable pedagogical implications for practitioners in the field of ELT. As for material designers, these findings may guide them in making informed decisions on how to design, select and sequence tasks within English language textbooks. Besides, the query into tasks in EFL textbooks would inform language teachers of the nature of tasks, the role of tasks in syllabus, task types as well as the cognitive/conceptual and performative/procedural demands tasks would pose on learners. In this way, teachers can make more informed decisions when designing, selecting, manipulating and sequencing the tasks to be implemented by a particular proficiency group to optimize teaching and learning.

The findings of the present study may also spur the researchers to further study task-based aspects in EFL textbooks such as tasks, task types, and task complexity and, therefore, explore their appropriateness for target proficiency groups as well as the opportunities they provide for the development of the target language. In this regard, it also bears the potential to contribute to the line of research which bridges TBLT and textbook design studies in the field of ELT.

Definitions of Key Terms

Within the present thesis, the terms that play a fundamental role in understanding this research are as follows:

Activity: Activity refers to both “the actual materials that constitute a task (i.e., the workplan)” and “the language use resulting from the performance of the task (i.e., the process)” (Ellis, 2018, p. 12; Ellis et al., 2019, p. 10).

Causal reasoning (dimension): It concerns whether a task requires "simple information transmission, versus reasoning about causal events and relationships between them" (Robinson, 2007, p. 18; Robinson, 2008, p. 7; Robinson & Gilabert, 2007, p. 165).

Closed tasks: Closed tasks denote the tasks designed to have only one or a limited number of outcomes, namely solutions (Ellis, 2012; Ellis et al., 2019).

Exercise: Exercise denotes “text-manipulation” activities undertaken to practice correct usage of the target language (Ellis, 2014, p. 107) and to facilitate “intentional learning” and the growth of “explicit knowledge” (Ellis, 2016, p. 213).

Few steps (dimension): It specifies whether one or few steps versus many steps are involved in a task (Robinson, 2007, 2008, Robinson & Gilabert, 2007).

Focused task: Focused tasks correspond to the tasks that aim at fostering communicative language use (Ellis, 2003) and orienting learners to processing predetermined, specific linguistic features (Ellis, 2009, 2012; Ellis et al., 2019).

Here-and-now task: They denote the tasks that enable task-takers to access the task stimuli during task implementation (Ellis, 2012; Ellis et al., 2019).

Here-and-now (dimension): It concerns whether tasks necessitate referring to "events happening now in a shared context (Here-and-Now) orient learner

attention to morphology for conveying tense and aspect in the present” or not (Robinson, 2010, p. 247; Robinson, 2015, pp. 15-16).

Independency of steps (dimension): It specifies if a task consists of activities to be completed in a strictly chained sequence following one after the other or not (Robinson, 2007, 2008, Robinson & Gilabert, 2007).

Input-based task: Input-based tasks correspond to listening or reading tasks (Ellis, 2009, 2011; Ellis et al., 2019). They are also called “receptive tasks” (Ellis et al., 2019, p. 156) as they do not require learners to produce language to complete a task.

Intentional reasoning (dimension): It concerns whether a task causes learners to reason about the intentions, beliefs and desires of people as well as the relationships between these or not (Robinson, 2007, 2008; Robinson & Gilabert, 2007).

Integrative task: Integrative tasks engage learners in using any skills from both input-based (i.e., listening and reading) and output-based tasks (i.e., speaking and writing) to complete a task.

Interlanguage: Interlanguage refers to “the linguistic system of learner language produced by adults when they attempt meaningful communication using a language they are in the process of learning” (Tarone, 2013, p. 2794).

L2: L2 refers to both second and foreign language in the present study as the foreign or second language distinction has weakened due to “(1) today’s exponentially expanded, Internet-based, border-spanning linguistic and cultural sharing and (2) the complexity of multicountry movements of individual immigrants and refugees” (Oxford, 2017, p. 28) and the concept is often used to denote both

second and foreign language learning situations in the recent literature (Byram & Hu, 2013; Oxford, 2017).

Open tasks: Open tasks refer to the tasks that have many possible outcomes, namely solutions (Ellis, 2011; Ellis et al., 2019).

Output-based task: Output-based tasks are the tasks that engage learners in language production through speaking or writing (Ellis, 2009, 2011, 2012, 2015, 2017, 2018; Ellis et al., 2019).

Pedagogic task: Pedagogic tasks are the tasks that seek interactional authenticity (Ellis, 2014, 2017, 2018; Ellis et al., 2019), which is achieved when the language produced during task implementation is similar to the one occurring in non-pedagogic conversation (Ellis, 2009).

Perspective-taking (dimension): It concerns if a task causes learners to adopt only one first-person perspective on an event (-perspective-taking) or assume several second and third-person perspectives (Robinson, 2007, 2008; Robinson & Gilabert, 2007) by using different pronominal references.

Planning time (dimension): It has to do with whether learners are provided with planning time or not (Robinson, 2007, 2008; Robinson & Gilabert, 2007).

Prior knowledge (dimension): It is related to the presence of background information provided to the task-takers (Robinson, 2007, 2008; Robinson & Gilabert, 2007).

Real-world (real-life) task: Real-world task - sometimes called “real-life task by Ellis (2020) - corresponds to the target tasks (Ellis et al., 2019), which language users usually perform in natural contexts outside school (Ellis, 2017). Such tasks have situational authenticity (Ellis et al., 2019) as they imitate real-life contexts (Ellis, 2009).

Resource directing variables: Resource directing variables refer to the factors that make “resource demands which can be met by using specific features of the language code” during task performance (Robinson, 2001b, p. 31).

Resource dispersing (depleting) variables: Resource dispersing variables denote the variables that make “demands on attention and working memory” while performing a task (Robinson, 2001b, p. 31).

Single task (dimension): It concerns if a task necessitates just one thing to complete or two or more things to be dealt with simultaneously (Robinson, 2007, 2008; Robinson & Gilabert, 2007).

Situational grammar exercise: Situational grammar exercise is a kind of exercise which contextualizes target forms through real or imaginary situations (Ellis, 1995).

Spatial reasoning (dimension): It concerns whether a task induces the employment of the linguistic forms to describe motion events or not (Robinson, 2011).

Task: It has been defined as a language learning activity that satisfies four criteria (Ellis & Shintani, 2014) as follows:

1. The primary focus should be on ‘meaning’ (i.e., learners should be mainly concerned with encoding and decoding messages not with focusing on linguistic form).
2. There should be some kind of ‘gap’ (i.e., a need to convey information, to express an opinion or to infer meaning).
3. Learners should largely rely on their own resources (linguistic and non-linguistic) in order to complete the activity. That is, learners are not ‘taught’ the language they will need to perform a task, although they may

be able to ‘borrow’ from the input the task provides to help them perform it.

4. There is a clearly defined outcome other than the use of language (i.e., the language serves as the means for achieving the outcome, not as an end in its own right). Thus, when performing a task, learners are not primarily concerned with using language correctly but rather with achieving the goal stipulated by the task. (pp. 135-136)

Task-based language teaching: TBLT has been operationalized as “contexts where tasks are the central unit of instruction: they ‘drive’ classroom activity, they define curriculum and syllabi, and they determine modes of assessment.” (Samuda & Bygate, 2008, p. 58).

Task-based syllabus: Task-based syllabus refers to the language syllabi in which tasks “function as the primary unit in selecting and sequencing course content” (Ellis et al., 2019, p. 179).

Task complexity: Task complexity has been described as “the result of the attentional, memory, reasoning, and other information processing demands imposed by the structure of the task on the language learner” (Robinson, 2001b, p. 29).

Task structure (dimension): Task structure is related to the presence or absence of a clear structure provided to support learners in determining what steps are involved in completing the task (Robinson, 2007, 2008, Robinson & Gilabert, 2007).

Task-supported syllabus: Task-supported syllabus denote the language syllabi in which tasks are used as supporting components, whereas other units of analysis such as “grammatical structures, lexis, topics, situations” constitute the primary organizing elements (Ellis et al., 2019, p. 179).

There-and-then (dimension): It concerns whether tasks necessitate referring to “events happening elsewhere in time and space (+There-and-Then)” or not (Robinson, 2010, p. 247; Robinson, 2015, pp. 15-16).

There-and-then tasks: There-and-then tasks refer to the tasks that do not allow task-takers to access task stimuli during task implementation (Ellis, 2012; Ellis et al., 2019).

Unfocused task: Unfocused tasks are the tasks that enable learners to use language in general communicatively (Ellis, 2009, 2012, 2012, 2015, 2018; Ellis et al., 2019; Ellis & Shintani, 2014).



CHAPTER 2: REVIEW OF RELATED LITERATURE

Introduction

This chapter presents an outline of the theoretical foundations of the present study that was conducted to investigate the tasks in the English textbooks in the Anatolian High school context in Turkey. With this aim, the first section of the literature review begins with the concept of task and provides a brief introduction to Task-based Language Teaching (TBLT). It continues with the discussion on the role of tasks in EFL textbooks. The following section explains task types. Delineating the previous proposals for sequencing tasks in language curricula in the field of English Language Teaching (ELT), the next section of the chapter explicates the issues of task difficulty and task complexity underpinned by the *Limited Attention Capacity* (LAC) and *Cognitive Hypothesis* (CH) by Skehan and Robinson, respectively. Finally, the last section of the chapter reviews some of the relevant studies previously conducted so as to justify the need for the present study.

Tasks and Task-based Language Teaching

Task, as one of the critical elements of language teaching, has long received attention from researchers in the field of ELT. Research on tasks and task-based learning has gained momentum since the 1980s with the preliminary studies by Breen (1989), Ellis (2003), Long (1985), Prabhu (1987), Robinson (2001a, 2001b), Skehan (1998) and Willis (1996).

Task-based research has become popular for several reasons. According to Gilabert et al. (2016), one of the motives to explore tasks is that implementing communicative tasks may direct learners' focus to language points and help them retain these language points, thus improving their IL systems. Besides, the

studies on tasks may be helpful in identifying the inherent characteristics of tasks that are likely to affect language processing to comprehend, produce and learn, thus yielding valuable pedagogical implications for designing materials and tasks (Gilabert et al., 2016). In turn, particular task types and designs may be purposefully chosen to generate opportunities for developing IL and communicative skills.

Task is considered the primary unit to determine and sequence the contents of a “task-based syllabus” (Ellis et al., 2019, p. 179). It can also be utilized to foster other types of syllabi designed based on different units of analysis such as “grammatical structures, lexis, topics, situations” (Ellis et al., 2019, p. 179). There are several reasons why tasks are suggested as a unit of syllabus. First, tasks create circumstances for “free” and “meaningful” use of the target languages, which is critical for learning (Willis, 1996, p. 30). Likewise, they are believed to be likely to create favourable circumstances for learners to practice L2 in meaningful ways by setting a framework for directing their attention to form (Révész, 2011). They also engage learners in “text-creation activities”, which bear the potential to foster the growth of implicit knowledge when they also involve a focus on form (Ellis, 2016, p. 215). According to skill-learning theory, such activities support learners in moving from the procedural to the automatic stage of language learning (Ellis, 2016, p. 216).

The incorporation of tasks into syllabus takes its roots from Communicative Language Teaching (CLT) and its successor, TBLT, which stands as an alternative to both traditional forms of language teaching (Ellis et al., 2019, p. 21) and CLT, which has drawn criticism for reaching a communicative content and procedure impasse (Samuda & Bygate, 2008).

In parallel with the growing interest in tasks and task-based approaches in language learning (Motteram & Thomas, 2010), TBLT has become a centre of

attention for practitioners and researchers in the field in the last quarter of the 20th century (Lai & Li, 2011). TBLT refers to “contexts where tasks are the central unit of instruction” and is considered as the elements that determine classroom activity, curriculum, syllabi and the assessment (Samuda & Bygate, 2008, p. 58). As an approach to curriculum (Gonzalez-Lloret & Ortega, 2014), TBLT has been developed to offer a theoretical base and practical implications for learning a second and foreign language (Lai & Li, 2011).

Although task-based language teaching denotes contexts in which tasks constitute the central unit of instruction, it should be noted that the notion of “task-based” has already extended to refer to all the contexts where tasks are utilized (Samuda & Bygate, 2008, p. 57) with the increasing popularity of tasks in English language teaching in various contexts (Butler et al., 2018).

Tasks and Task Characteristics

To understand the notion of task, it is crucial to discuss what *task*, *non-task* (i.e., *exercise*), and *activity* denote, which are often used interchangeably in ELT. The operationalization of these terms will set one of the conceptual frameworks of the present study and aid our understanding of tasks. With this purpose in mind, the following sections are devoted to explicating each concept. First, the concept of *activity* is explained since it includes both tasks and non-tasks. Please note that the notion of *non-task* refers to *exercise* in the present study, as in Ellis (2003). Secondly, the concept of *task* is introduced. Finally, the concept of *non-task* is discussed with specific regard to the differences between tasks and non-tasks (i.e., exercises).

Activity

The first concept that needs to be distinguished from task is *activity*. Upon surveying the literature on task-based learning, it has been observed that the concept of activity is an umbrella term involving both tasks and non-tasks. Task is considered a type of activity that meets specific criteria of taskness (Ellis, 2003; Prabhu, 1987; Skehan, 1996; Willis, 1996). When the concept of activity is used to refer to tasks, it may denote different stages of tasks. That is, activity applies to both “the actual materials that constitute a task (i.e., the workplan)” and “the language use resulting from the performance of the task (i.e., the process)” (Ellis, 2018, p. 12; Ellis et al., 2019, p. 10). Ellis prefers to use this concept to refer to both “materials” including “texts, exercises and tasks” as well as the “procedures” to utilize them as “metalinguistic explanations and corrective feedback” (2012, pp. 16-17).

The conceptualization of activities as “materials” indicates that task is taken as a workplan that involves both tasks and non-tasks, while the “procedures” refers to task-as-process (Ellis, 2018, p. 12; Ellis et al., 2019, p. 10). In the present study, the concept of activity refers to both task as material, including tasks and non-tasks, as well as task as procedure as conceptualized by Ellis (2003, 2012, 2018).

Task

There has not been an agreement on defining the concept of *task* in the ELT literature. Several definitions have been proposed (Breen, 1989; Ellis, 2003, 2009; Nunan, 2004; Prabhu, 1987; Willis, 1996) for task. Another concept, “task-like”, has also been conceptualized differently in previous studies (Dao & Newton, 2021, p. 106; Jaruteerapan, 2020, p. 9; Terekhova, 2017, p. 37).

One of the earliest definitions of task was made by Prabhu (1987, pp. 24-27), according to whom it is a “meaning-focused activity in which learners are occupied

with understanding, extending (e.g. through reasoning), or conveying meaning, and cope with language forms as demanded by that process” and “attention to language forms is thus not intentional but incidental to perceiving, expressing, and organizing meaning” to arrive at a task outcome.

Another definition is offered by Breen (1989, p. 157), who interprets the concept of task as “a structured plan for the provision of opportunities for the refinement of knowledge and capabilities entailed in a new language and its use during communication” and describes it as a kind of a workplan with its specific objectives, contents and working procedures.

Skehan (1996, p. 38) defines task as “an activity in which meaning is primary, there is a relationship to the real world, task completion has some priority, and the assessment of task performance is in terms of task outcome”.

Another researcher, Willis (1996, p. 23), defines task as an activity “where the target language is used by the learner for a communicative purpose (goal) in order to achieve an outcome”.

Nunan (2004, p. 4) offers a definition of what he calls a “pedagogical task” as “a piece of classroom work that involves learners in comprehending, manipulating, producing or interacting in the target language” with their attention focused on meaning, arguing that it must “have a sense of completeness, being able to stand alone as a communicative act in its own right with a beginning, a middle and an end”.

A closer look at task definitions reveals that the most frequently emphasized aspects of tasks include focus on *meaning* (Nunan, 2004; Prabhu, 1983; Skehan, 1996; Willis, 1996), structured *workplan* (Breen, 1989; Nunan, 2004), communicative *purpose* (Breen, 1989; Nunan, 2004; Willis, 1996) and task *outcome*

(Prabhu, 1987; Skehan, 1996; Willis, 1996) that signify that a task is completed (Nunan, 2004; Skehan, 1996). It can be inferred from all of the task operationalisations aforementioned that language is viewed as a practical means catering to the task's goals.

Drawing on previous definitions offered, Ellis (2003) proposes a definition for task in his seminal work, *Task-based Language Learning and Teaching*:

a workplan that requires learners to process language pragmatically in order to achieve an outcome that can be evaluated in terms of whether the correct or appropriate propositional content has been conveyed. To this end, it requires them to give primary attention to meaning and to make use of their own linguistic resources, although the design of the task may predispose them to choose particular forms. A task is intended to result in language use that bears a resemblance, direct or indirect, to the way language is used in the real world. Like other language activities, a task can engage productive or receptive, and oral or written skills, and also various cognitive processes. (p. 16)

Based on this definition, Ellis (2003) identifies six criterial characteristics of a task, which are:

1. Task is a workplan.
2. A task involves primary focus on meaning.
3. A task involves real-world processes of language use.
4. A task can involve any of the four language skills.
5. A task engages cognitive processes.
6. A task has a clearly defined communicative outcome. (pp. 9-10)

In line with the previous definitions for task in TBLT, Ellis (2009, 2011, 2012) revises his earlier criteria for an activity to be considered as task, giving them the final version in 2014 (Ellis & Shintani) as follows:

1. The primary focus should be on ‘meaning’ (i.e. learners should be mainly concerned with encoding and decoding messages not with focusing on linguistic form).
2. There should be some kind of ‘gap’ (i.e. a need to convey information, to express an opinion or to infer meaning).
3. Learners should largely rely on their own resources (linguistic and non-linguistic) in order to complete the activity. That is, learners are not ‘taught’ the language they will need to perform a task, although they may be able to ‘borrow’ from the input the task provides to help them perform it.
4. There is a clearly defined outcome other than the use of language (i.e. the language serves as the means for achieving the outcome, not as an end in its own right). Thus, when performing a task, learners are not primarily concerned with using language correctly but rather with achieving the goal stipulated by the task. (pp. 135-136)

The first question of the present study investigates to what extent the activities in EFL textbooks possess task characteristics. The study draws on the four criteria mentioned above to distinguish tasks from non-tasks (i.e., exercises). Although the definitions by Ellis (2003, 2009) and Ellis and Shintani (2014) are among the most commonly cited ones for their clearer representations of *task*, the notion remains “somewhat fuzzy” (Richards, 2005, p. 31). For this reason, each criterion will be explained in detail in the following sections by referring to Ellis’s

conceptualization of the term in his studies up to the present (Ellis, 2000, 2003, 2005, 2006, 2009, 2011, 2012, 2016, 2018; Ellis et al. 2019; Ellis & Shintani, 2014).

Criterion 1

The first criterion, “the primary focus should be on meaning” (Ellis, 2009, p. 223; Ellis & Shintani, 2014, p. 135), means that the workplan requires language learners to comprehend and produce messages for communication (Ellis, 2018; Ellis et al., 2019). It denotes the activities that engage language learners in employing the target language both pragmatically and indexically to communicate their meaning intentions (Ellis & Shintani, 2014). While “semantic meaning” implied in the criterion is defined as “the meanings of lexical items or specific grammatical structures”, “pragmatic meaning” is associated with “the highly contextualized meanings that arise in acts of communication” (Ellis, 2005, p. 12) for such functions as describing, requesting, apologizing and so on (Ellis & Shintani, 2014) in natural contexts of use (Ellis, 2009). Therefore, for an activity to meet the first criterion, it must engage learners in using language both pragmatically and indexically for communicative purposes (Ellis & Shintani, 2014).

No matter whether a task-based or task-supported approach to language teaching is adopted, it is recommended that learners are engaged in encoding and decoding both types of meaning, especially pragmatic meaning, since it is considered critical for learning languages (Ellis, 2005). Paying attention to semantic or pragmatic meaning differs in that language is treated as an object by learners in the first one, whereas it is regarded as a tool for communication by communicators in the second one (Ellis, 2005).

While learners focus on meaning, they act as a “language user” not as a “language learner” (Ellis, 2003, p. 3; Erlam, 2016, p. 288) since they are somehow

engaged in identical communicative processes with the people performing real-world activities (Ellis, 2003).

Although language learners are expected to act as “communicators” instead of “learners” so that tasks function as planned, it is also acknowledged that this expectation is unrealistic as students are likely to keep their learner roles during task implementation (Ellis & Shintani, 2014, p. 143). Students focus primarily on message conveyance and secondarily on language forms while performing tasks and as their focus switches between meaning and form (i.e., code) on and off during task implementation (Lambert, 2018), the roles of the students also shift from and to language learner and user temporarily (Ellis, 2003). Hence it is at least desired that students switch between these roles while trying to accomplish the tasks (Ellis & Shintani, 2014).

One example of the application of this criterion is the *Spot-the-Difference Task* in which learners are expected to compare two pictures and pay attention not only to the physical features and the spatial location of the items and, thus, semantic meaning; but also to producing language so that they can link what they encode or decode to the given images and their partner(s)’ utterances (i.e., pragmatic meaning) as well as to the emerging communication problems (Ellis & Shintani, 2014).

Criterion 2

The second criterion, “there should be some kind of ‘gap’” (Ellis, 2009, p. 223; Ellis & Shintani, 2014, p. 135), underlines the need for communicating in the target language. This need functions as a motivation factor to communicate so as to close the gap (Ellis, 2009).

The concept of *gap* in the task definition by Ellis and Shintani (2014) draws on the study by Prabhu (1987, pp. 46-47) who proposes three types of gaps as

“information gap”, “opinion gap”, and “reasoning gap”. The first one is *information gap* tasks that involve information provided in a split fashion to the task-takers who need to exchange their piece to close the gap and accomplish the task goals (Ellis, 2003).

Information gap tasks can be one-way or two-way depending on who has and supplies information. While one learner has the information as a whole and communicates it to the other learner in one-way tasks, each participant possesses one piece of information that their partner lacks in two-way tasks (Ellis & Shintani, 2014). While the task-takers act as a speaker or listener in one way tasks, they take both roles in two-way tasks (Ellis & Shintani, 2014). An example of a two-way task is jigsaw task (Ellis & Shintani, 2014).

In the *opinion gap* tasks, learners need to offer their own ideas beyond the information provided in the task input (Ellis, 2003). In this case, the exchange of information is not obligatory (Ellis, 2003). Instead, they typically engage task-takers in an exchange of opinions on issues open to debate (Ellis & Shintani, 2014). Learners, therefore, are required to identify and state “preferences”, “feelings”, or “attitudes” to address given situations (Ellis et al., 2019, p. 183). For instance, participants may share information about four patients (two-way information gap) and then decide which one deserves the heart transplant most by expressing their own opinions (opinion gap) as in the *Heart Transplant Task* (Ellis & Shintani, 2014).

Finally, *reasoning-gap* tasks necessitate drawing new information through inference from the already provided piece of information (Ellis & Shintani, 2014). They engage task-takers in cognitive processes of synthesis, deduction (Ellis & Shintani, 2014), “inference, calculation, collation, application or identification of relationships/patterns” (Ellis et al., 2019, p. 183).

Criterion 3

The third criterion, “learners should largely rely on their own resources”, means that there is no dictation of particular linguistic features to be used during task implementation (Ellis, 2009, p. 223; Ellis & Shintani, 2014, p. 135), but learners may borrow from the task input if they need (Ellis, 2009, 2018; Ellis & Shintani, 2014). This requires that learners do not have access to model language, nor are they instructed or suggested to focus on particular language forms (Lambert, 2018). Besides, there should be no instruction on how to use language or perform interaction, even if a task might naturally induce the employment of specific linguistic forms (Ellis, 2003). For instance, some activities that require learners to answer multiple choice questions or fill in the blanks do not give them the chance to have a choice over linguistic resources to utilize and, therefore, fail to qualify as a task (Ellis, 2003).

It is expected that learners utilize their own resources such as gestures, facial expressions, L1, L2 (Ellis, 2018; Ellis et al., 2019) as well as “prior knowledge of vocabulary, context, world knowledge” (Jaruteerapan, 2020, p. 71) for comprehension or production. In this way, it is up to learners which social acts to perform, such as making a request, refusing and persuading (Ellis, 2003). The number of linguistic resources utilized to produce the task outcome may change based on the nature of the task and the proficiency levels of the learners (Ellis, 2014).

This criterion concerns to what extent learners improve their ability to access their (non)linguistic resources and utilize them to satisfy task and conceptual demands (Lambert, 2018). When learners use their own resources, they are believed to work out new strategies for communication in L2 (Lambert, 2018).

Criterion 4

The criterion, “there is a clearly defined outcome other than the use of language” (Ellis, 2009, p. 223; Ellis & Shintani, 2014, p. 135), considers that a task has a specific, non-linguistic outcome, the achievement of which signals its completion (Ellis, 2003, 2020; Erlam, 2016). It means that language serves as a tool for the task outcome instead of being an end itself (Erlam, 2016). It also indicates that practising language is not enough for an activity to qualify as a task.

In this sense, the last criterion agrees with Nunan’s (1989, p. 10) definition of task, which bears “a sense of completeness” to function as a “communicative act” by itself. The absence of a specific tangible or intellectual outcome, therefore, necessitates excluding free, unstructured conversation activities from the task category (Cobb, 2010).

A task outcome can be verbal (linguistic) or nonverbal (nonlinguistic) depending on whether it requires learners to comprehend or produce language. On the one hand, it can be only non-verbal in the *input-based* tasks, in which learners may engage in activities such as sequencing pictures, completing a map and creating a model. In the output-based tasks, it can be verbal and/or nonverbal (Ellis, 2018) such as creating a “story”, making a “list of differences” (Ellis, 2003, p. 154), solving a problem, making a decision, and identifying a someone (Cobb & Lovick, 2007).

To achieve the outcomes, learners may engage in several cognitive processes including “selecting, classifying, ordering, reasoning and evaluating information” (Ellis, 2003, p. 10). Even if the primary reason for engaging learners in tasks is to make them practice language by going through cognitive and linguistic processes to promote their learning, tasks should be designed in a way to focus their attention on achieving the task outcome (Ellis, 2003). In addition, the cognitive processes learners

employ during task implementation should not be confused with task outcome as they basically serve to accomplish the predefined communicative outcome. The outcome of a task, therefore, should be judged independently from the mere use of the target language or the performance of the activity itself; instead, it should be evaluated based on its content (Ellis, 2003, p. 8-15).

Task *outcome* should also be distinguished from the *aim* of a task (Ellis, 2003), which is a common mistake made (Jaruteerapan, 2020). As task workplan represents merely the intended outcome of a task and cannot guarantee that task implementation will yield the predicted language use, it is probable that task takers accomplish the outcome of a task while failing its aim (Ellis, 2003).

Interpretation of Task Criteria by Erlam (2016) and Jaruteerapan (2020)

The task criteria by Ellis (2003, 2009, 2018) have been taken as a reference to find out the extent to which activities possess the task characteristics in several studies (Barrot, 2019; Butler et al., 2018; Dao & Newton, 2021; Terekhova, 2017). However, they have drawn criticism for being somewhat ambiguous and two studies were devoted to the explanation of these criteria (Erlam, 2016; Jaruteerapan, 2020).

The first study by Erlam (2016, p. 288) attempted to explicate each criterion with two conditions phrased in questions (Table 1). A checklist was generated to determine whether an activity can be considered as a task or non-task.

Table 1

Questions Relating to Coding of Criteria by Erlam

Questions		Required answer
1	The primary focus should be on meaning	
	Does the learner function as a language user and not a language learner?	Yes
	Is the learner primarily concerned with encoding and decoding messages, not with focusing on linguistic form?	Yes

Table 1 (cont'd)*Questions Relating to Coding of Criteria by Erlam*

	Questions	Required answer
2	There should be some kind of gap	
	Is this gap closed as a result of the communication that takes place?	Yes
	As a result of the communication does the learner find out something they didn't know?	Yes
3	Learners should have to rely on their own resources (linguistic and nonlinguistic)	
	So that they can do the task, are the learners 'taught' the language they will need?	No
	Does the task allow learners to automatize/use language they have already been taught on a previous occasion?	Yes
4	There is a clearly defined outcome other than the use of language	
	Does the language serve as a means for achieving the outcome rather than as an end in its own right?	Yes
	Does achieving the outcome determine when the task is completed?	Yes

Note. From “‘I’m still not Sure What a Task is’: Teachers Designing Language Tasks”, by R. Erlam, 2016, *Language Teaching Research*, 20(3), p. 288 (<https://doi.org/10.1177/1362168814566087>). Copyright 2016 by Language Teaching Research.

For the first criterion (focus on meaning), a learner is expected to (1) function as a “language user” rather than a “language learner” and (2) be primarily concerned with encoding and decoding messages rather than linguistic form (Erlam, 2016, p. 288). For the second criterion (gap), it is expected that (1) the gap is closed as a result of communication taking place and (2) learner finds out something they did not know. Erlam (2016) conceptualises the *gap* in Ellis’s framework as the existence of an audience who read or listen for a purpose. This interpretation of gap is also in line with how Willis (1996) conceptualises task objectives, who is of the opinion that there must be a purpose to read or take notes, as there must be an audience to speak to or write for so that meaningful interaction occurs. Thirdly, learners (1) should not

be taught the language they will need and (2) are allowed to automatize/use language already being taught on a previous occasion for the third criterion (using their own resources). Finally, it is expected that (1) language functions as a means to achieve the outcome instead of being an end in its own right and (2) achieving the outcome determines the completion of a task.

Commenting that Ellis's (2009) task criteria might be misinterpreted to denote solely productive language when Erlam's (2016) interpretation is taken into account, Jaruteerapan (2020) revised Erlam's interpretation with the aim of emphasising both receptive and productive forms of communication and developed a new coding framework for tasks (pp. 71-72) (Table 2).

Table 2

Coding Framework for Identifying Tasks by Jaruteerapan

Criteria	Questions	Required answer
1. A primary focus is on meaning. (C1)	1.1 Do the learners mainly use the language for a communicative purpose, not just learn the language (e.g., forms and functions) intentionally?	Yes
	1.2 Is the learner primarily concerned with producing and/or comprehending the messages (semantic and pragmatic meaning), not with focusing on linguistic form?	Yes
2. There should be some kind of gap. (C2)	2.1 Does the activity have a gap that requires a transfer of given message from one person to another, or one form to another (written to oral) or one place to another (texts to table) in order to close it?	Yes
	2.2 Is the gap closed as a result of learners' participation in communication (i.e., as either as a speaker/writer or listener/reader)?	Yes

Table 2 (cont'd)*Coding Framework for Identifying Tasks by Jaruteerapan*

3. Learners rely on their own resources. (C3)	3.1 Are the learners not “taught” * the language they will need to perform the activity? * <i>Being taught the language does not include the teachers providing some linguistic starting point such as key vocabulary, an input for production tasks (priming)</i>	Yes
	3.2 Does the activity allow the learners to use the language they already knew to complete the activity?	Yes
4. There is a clearly defined communicative outcome. (C4)	4.1 Are the learners primarily concerned with achieving the goal stipulated by the activity, rather than using language forms correctly?	Yes
	4.2 Is there an outcome that results from completing an activity that works towards a communicative goal, other than the display of linguistic knowledge?	Yes

Note. Adapted from *The Emerging Understandings and Practices of Task-Based Language Teaching (TBLT)* by Thai EFL Student Teachers (pp. 71-72), by P. Jaruteerapan, 2020, [Unpublished Doctoral Dissertation] Victoria University of Wellington. Copyright 2020 by Victoria University of Wellington.

As given in Table 2, Jaruteerapan (2020) develops two conditions that should be satisfied for an activity to be considered a task, explicating Ellis’s (2009) first criterion. First, learners utilise language mainly for a communicative purpose, not only to learn it. Second, they focus mainly on producing or comprehending messages, not linguistic form.

Comparing the questions developed by Erlam (2016) and Jaruteerapan (2020) for the first task criterion, it is observed that Jaruteerapan (2020) keeps the original question by Erlam (2016) and adds “mainly” (Jaruteerapan, 2020, p. 71) for clarity purposes when developing the first question of the first criterion. The words “language user” and “language learner” (Erlam, 2016, p. 288) are not preferred as

learners may switch back and forth between these roles, focusing on form and meaning on and off. The words in the second question of the first criterion, “encoding” and “decoding” (Erlam, 2016, p. 288), are also replaced with “comprehending” and “producing” for clarity (Jaruteerapan, 2020, p. 71).

For Jaruteerapan (2020, p. 71), for the second criterion to be satisfied, (1) a task should have a gap that necessitates “a transfer of given message from one person to another, or one form to another (written to oral) or one place to another (texts to table) in order to close it” and (2) the gap should be closed when learners communicate.

Table 2 shows that the questions by Erlam (2016) for the second criterion are also revised. Initially, the first question is modified to explicate what the notion of gap involves in the first question. Then, the examples of communication are specified in the second question. The modification of the questions for the second criterion by Erlam (2016) is helpful as it makes it clear that a transfer of a message can occur in different ways, and communication can be receptive as well as productive.

As for the third criterion, Jaruteerapan (2020) explicates it with two conditions to be met. First, learners are not taught the language required to accomplish the task. Second, they are allowed to use the language they already know. It means that learners are free to make their own selection of any linguistic or non-linguistic resources available when performing the task given.

When two frameworks are compared, it is observed that the first question in Erlam’s (2016) coding of tasks is kept in Jaruteerapan’s (2020) coding framework for tasks. However, the negative version of the question is used in the revised version and the meaning of “taught” (Erlam, 2016, p. 288) is explained with an additional

note. The second question is modified to avoid confusion. The expression “language they have already been taught” is replaced by “language they already knew” to underline learners' current resources (Jaruteerapan, 2020, p. 320).

In the final version of the questions for the last criterion, it is required that (1) learners mainly focus on achieving the task goal instead of correct use of language forms and (2) the task outcome serves a communicative goal, not the display of linguistic knowledge. It can be commented that the questions for the fourth criterion are modified to emphasise “a clearly defined outcome” and “the display of language” (Jaruteerapan, 2020, p. 320).

Jaruteerapan (2020, p.72) specifies that a communicative outcome can be an “observable result” (e.g., “a completed note or table, a route drawn on a map, a list of differences between 2 pictures, a survey, a drawing, a completed reorder of jumbled items, etc.”), “an abstract idea” (e.g., “approved decision, a plan for problem solving, a consensus etc.”) or a “physical reaction” (e.g., “outcomes from listen-and-do activity”).

When the coding frameworks developed by Erlam (2016) and Jaruteerapan (2020) for task criteria are analysed, it seems that the one by Jaruteerapan (2020) provides a more detailed and comprehensive explanation of the task criteria by Ellis (2009) as it attempts to avoid confusion that might arise due to wording in the questions developed for each criterion by Erlam (2016) and to provide additional notes and examples. Therefore, the task criteria coding framework by Jaruteerapan (2020) would serve as a valuable source when analysing language activities and determining which activities qualify for a task.

Non-Task (Exercise)

Although the characteristics of tasks have been widely discussed in the ELT literature, the concept is often confused with *exercise*, which is considered *non-task* in the present study. Ellis distinguishes *task* from *non-task* (i.e., *exercise*) (Ellis, 2009, 2006; Ellis & Shintani, 2014). Ellis considers traditional language learning activities as exercise, which include “situational grammar activities” (2009, p. 227), “repetition, substitution, gapped sentences, jumbled sentences, transformation drills, and dialogues” (2006, p. 96). In line with the conceptualizations of Ellis (2003), *non-task* activities (i.e., the activities that do not qualify for a task) are called *exercise*.

Exercises (i.e., *non-tasks*) primarily function to help learners practise specific language points (Ellis & Shintani, 2014). As opposed to tasks, exercises do not prioritize meaning, present a gap to be closed, require learners to use their own linguistic resources or yield an outcome distinguishable from the mere practice of the target language (Ellis & Shintani, 2014). On the contrary, they tend to inform learners of the particular language points to work on and require them to manipulate these language points correctly (Ellis & Shintani, 2014).

Another term Ellis uses to distinguish tasks from non-tasks is situational grammar exercise”, which refers to a specific type of “an exercise that has been designed to provide contextualised practice of a specific linguistic feature” (Ellis, 2003, p. 141). Situational grammar exercises may meet some of the four criteria proposed, especially criteria 2 and 3 (Ellis, 2009, 2012). However, such exercises generally fail to satisfy criteria 1 and 4, as they serve to practice predetermined language features rather than creating communicative contexts for using language as a means (Ellis & Shintani, 2014) and do not focus on meaning or producing a communicative outcome but the performance of correct language (Ellis, 2009, 2012).

Exercises (i.e., non-tasks) entail “text-manipulation” activities such as “filling in gaps in sentences; performing a scripted dialogue; substituting words in a model sentence/dialogue” to practice correct usage of the target language (Ellis, 2018, p. 107) and facilitate “intentional learning” and the growth of “explicit knowledge” (Ellis, 2016, p. 213), while tasks engage learners in “text-creation” as they are expected to use their own means to decode input or encode messages in terms of not only semantic but also pragmatic meaning to close a *gap* to produce a communicative outcome to accomplish a task (Ellis, 2018, p. 265).

Ellis (2016) observes that ELT textbooks consist of much fewer independent text-creation workplans (tasks) compared to text manipulation ones (exercises) and that text-creation workplans do not follow text-manipulation workplans. Ellis comments that those coursebooks consider target language not as a *tool* but mainly as an *object*. This situation, according to him, decreases the chances that the coursebooks support “user-based learning” or the development of “implicit knowledge” or “automized declarative knowledge”, which are necessary for communication in the target language (p. 216).

Tasks versus Non-tasks

It should be noted that although the distinction between task and non-task (i.e., exercise) is made in the ELT literature, it should not be concluded that the latter is pedagogically insignificant (Ellis, 2009, 2013). It is emphasised that both tasks and exercises are valuable for supporting various aspects of language learning (Ellis, 2013, 2018). The ultimate aim of each is to learn the target language, even if the means they choose differ (Ellis, 2003, p. 3). The emphasis on tasks and TBLT is made not to suggest that they should replace traditional forms of teaching, but that

TBLT can be incorporated into language teaching alongside these conventional forms of teaching.

The positions of activities that are form-focused (i.e., exercise) or meaning-focused (i.e., task) are commonly considered as a continuum rather than a dichotomy (Ellis & Shintani, 2014; Estaire & Zanon, 1994; Littlewood, 2004). The shift from the focus on form to meaning is interpreted to move an activity from “non-communicative learning” to “authentic communication” (Littlewood, 2004, p. 322), from “exercise” to “task” (Ellis, 2000, p. 197) and from “enabling tasks” to “communicative tasks” (Esaire & Zanon, 1994, pp. 13-15).

It is not a straightforward process to decide whether an activity qualifies as a task or an exercise as it may display the characteristics of both (Ellis, 2003). Since the distinction between task and exercise is not watertight, it can be considered as a continuum (Ellis, 2018; Ellis et al., 2019; Jaruteerapan, 2020; Littlewood, 2007), from activities possessing all the characteristics of a task to the activities that bear some task characteristics and finally to the ones exhibiting none. In this case, activities can be more or less task-like (Ellis, 2018; Ellis et al., 2019).

Task as Unit of Syllabus

To be able to discuss the role of tasks in a syllabus, it is essential to mention what the notions of curriculum and syllabus denote. *Curriculum* pertains to “the planning, implementation, evaluation, management, and administration of education programmes”, whereas *syllabus* is concerned with “what” of the educational programme and the scope of syllabus restricted to “the selection and grading of content” (Nunan, 1988, pp. 6-8) as a part of a “wider and more general curriculum” (Byram & Hu, 2013, p. 679). While syllabus reflects a particular curriculum, syllabus itself can be implicitly reflected in a textbook (Byram & Hu, 2013). Broadly

speaking, a comprehensive analysis of a language textbook in terms of tasks can provide information about the role of tasks in a syllabus.

Breen suggests that tasks are evaluated in three phases to develop tasks that better serve their purpose: “task-as-workplan”, “task-as-process”, and “task outcomes” (1989, p. 188). Referring to Breen’s (1989) classification of tasks, Ellis et al. (2019) argue that this classification is valuable for understanding the distinction between syllabus and methodology. While “task-as-workplan” informs syllabus design, “task-as-process” informs the methodology (Ellis et al., 2019, p. 196). Distinguishing the “design” and “implementation” phases, Ellis (2017, p. 509) contends that task should be identified merely based on the design features of a workplan. Since syllabus functions as a workplan that guides language teachers to help learners develop language competencies instead of teaching them to implement tasks (Ellis et al., 2019), task-as-workplan (i.e., design materials) is suggested to be the starting point that informs course design, testing and research (Ellis, 2018; Ellis et al., 2019). Ellis (2018) maintains that it would not be right to define a task as a process as this stage is somewhat unpredictable, considering that task-as-process is determined by not only how a task is designed but also how it is performed.

There is always the possibility that teachers make changes to tasks, subvert or even ignore them during teaching (Ellis et al., 2019). Besides, task-takers are likely to handle tasks from different perspectives than their designers (Ellis et al., 2019). The study by Berben, Van den Branden, and Van Gorp (2007, p. 56), for instance, demonstrate how many different guises a task can take, such as “intended task”, “expected task”, “performed task”, and “assessed task” in different stages of learning, commenting that a task cannot be taken as a fixed identity. Hence, it is

acknowledged that the workplan may not result in the intended outcome when implemented (Ellis, 2003).

Researchers in the *Sociocultural Theory* distinguish *task*, “the actual task materials that are given to the learners,” from *activity*, namely “the actual performance of the task”, and argue that an activity cannot be determined by a task since the performance of the task-takers draw on their unique goals, motives and strategies (Ellis, 2012, p. 15). Besides, *Activity Theory* claims that what is considered a “fixed task” can become rather variable when it is implemented as an activity shaped by several factors, including the personal objectives, level of knowledge and willingness of the task-takers (Coughlan & Duff, 1994, p. 174). Though acknowledging the difference between *task* (i.e., actual task materials) and *activity* (i.e., task performance) as defined by sociocultural theorists, Ellis (2012) argues that task and activity are not entirely distinct from each other, but there is a link between them. The manipulation of some task procedures, for instance, may result in intended language performance (Ellis, 2003).

The arguments by Ellis (2009, 2012) do not entirely conflict with the ones put forward by *Sociocultural Theory* and *Activity Theory*. Both maintain that there is always a possibility that how a task is implemented can differ from how it is designed. Nevertheless, focusing only on task-as-process and ignoring task-as-workplan would be an oversimplification of the role of the task as a design material as the effect of task as workplan on task performance has been revealed in previous research several times (Adams & Alwi, 2014; Arslanyılmaz & Pedersen, 2010; Awwad, 2017; Awwad et al., 2017; Cho, 2011; Choong, 2014; Donate, 2018; Gilabert, 2007; Gilabert et al., 2009; Hardy & Moore, 2004; Hsu, 2012; Larson, 2020; Lee, 2018; Luo, 2007; Moradi & Farvardin, 2019; Révész, 2011; Riccardi,

2014; Sasayama, 2015; Tabari, 2018; Vo, 2018; Yıldız, 2017). Therefore, taking task-as-workplan as a reference when examining EFL textbooks could be a useful starting point to predict what demands tasks can potentially impose on learners and make informed decisions on the design, selection, and sequencing of tasks in the textbooks.

Task-based versus Task-supported Syllabus

The goals and scope of syllabi determine the approach adopted for syllabus design (Ellis et al., 2019). Two approaches commonly cited to explicate the place of tasks in a language syllabus are *task-based syllabus* and *task-supported syllabus*. Tasks can be the main component of a syllabus as in task-based syllabus or utilized to foster other types of syllabi designed based on different units of analysis such as “grammatical structures, lexis, topics, situations” (Ellis et al., 2019, p. 179).

Task-based language teaching necessitates an instructional programme identified in terms of tasks and a syllabus comprising *unfocused* tasks (Ellis, 2009), namely the tasks that enable learners to use language in general communicatively (Ellis, 2009, 2012, 2012, 2015, 2018; Ellis et al., 2019; Ellis & Shintani, 2014). In this approach, task is considered as the primary unit to determine and sequence the contents of a *task-based syllabus* (Ellis et al., 2019) and a critical component for facilitating key processes of language acquisition (Samuda & Bygate, 2008).

As opposed to task-based language teaching, *task-supported language teaching* requires a structural syllabus, where tasks may still have a crucial role but are not the primary component that defines the organization of syllabus or assessment procedures (Révész, n.d.; Samuda & Bygate, 2008). That is, tasks in task-supported syllabus constitute only one of the elements (e.g., exercises, focused

practices) within the overall instruction program to accomplish different pedagogic purposes (Samuda & Bygate, 2008).

Although the distinction between task-based and task-supported language teaching is emphasized in the ELT literature, these concepts are generally confused, and “task-based” is often used to refer to any context in which tasks are used (Samuda & Bygate, 2008, p. 57). It has also been observed that tasks and TBLT have different connotations to different stakeholders in language teaching (Long, 2016). For instance, traditional exercises with a linguistic focus are introduced as tasks within some commercial course books and literature on pedagogy (Long, 2016). Therefore, one needs to analyse how instructional programs, syllabi or language textbooks are structured to make accurate judgments about whether they are task-based or task-supported.

Task Cycle

Several researchers conceptualized a task-based lesson as a three-stage process, which consists of the phases before, during, and after the main task (Ellis, 2006; Willis, 1996). Willis (1996, p. 38) names these phases “pre-task”, “task-cycle”, and “language focus”. In this format of a task-based lesson, the first phase, “pre-task”, functions to introduce the topic and the task to prepare learners for the main task (p. 38). The second phase, “task-cycle”, includes the task, planning, and report stages (p. 38). While learners implement a task during the “task” phase, they get prepared to report their task implementation process in the “planning” phase and present these reports in the “report” phase (p. 38). The last phase, “language focus”, on the other hand, concerns the analysis and the practice of specific linguistic features (p. 38).

Ellis names (2006, p. 20) the three phases of a task-based lesson as “pre-task”, “during task”, and “post-task”. These stages have similar functions to the framework Willis (1996) proposes.

The first stage, “pre-task”, serves to frame the activity, give task-takers time to plan their task performance (i.e., strategic planning), and presents a model or similar activity to the main task (Ellis, 2006, p. 20). Additionally, non-task preparation activities (i.e., exercise) can be offered at this stage (Willis, 1996); prediction, cooperative dictionary search, word-meaning matching (Newton, 2001). “During task” phase involves the main task performance and can be manipulated in terms of task performance or process options (Ellis, 2006, p. 20). Finally, the “post-task” stage involves follow-up activities (Ellis et al., 2019, p. 209) and engages task-takers in preparing a report of their task experience, performing consciousness-raising tasks (focusing on form) and repeating the same task (Ellis, 2006, p. 20). Of the three phases, a task-based lesson should minimally include one “during task” (i.e., main task) stage (Ellis, 2006, p. 20); however, pre-task and post-task stages are functional for ensuring task implementation works towards language development (Ellis, 2003).

Task Types and Sequencing of Tasks

The selection of the task types is among the most critical issues in TBLT. Research on the classification and sequence of tasks is encouraged to inform how to design task-based syllabi (Long, 2015). It is believed that classifying tasks can be practical for several reasons, such as to diversify the task types in syllabus, to provide the right kinds of tasks for particular needs and preferences, and to evaluate tasks to discover which ones work for specific learner profile (Ellis, 2003). The classification of the task types is also essential for selecting appropriate tasks to

incorporate into syllabus as the types of the tasks presented in textbooks are considered one of their quality indicators (Huda & Syafei, 2020).

Several task typologies have been offered in the ELT literature (Ellis, 2003; Long, 1985; Nunan, 2004; Pica, Kanagy, & Falodun, 1993; Prabhu, 1987; Willis, 1996). One of the earliest typologies belongs to Long (1985, p. 91), who mentions two types of tasks: “target” and “pedagogic” tasks. In another typology by Nunan (1989, p. 40), communicative tasks are categorized as “real-world” and “pedagogic” tasks based on the real-world or pedagogic rationale for their design. Another typology is offered by the same researcher (Nunan, 2004, p. 59) based on the strategies underpinning tasks, which are “cognitive, interpersonal, linguistic, affective and creative” strategies.

Surveying the literature on task typologies, Ellis (2003, p. 211) groups the approaches to task classification under four categories: “pedagogic”, “rhetorical”, “cognitive” and “psycholinguistic”. One of the most prominent figures in pedagogic typology is Willis (1996, pp. 26-27), who determines six categories: “listing”, “ordering and sorting”, “comparing”, “problem-solving”, “sharing personal experiences”, and “creative tasks”. Secondly, rhetorical classification depends on the structure and linguistic features of discourse domains and is concerned with genres such as “narrative, instructions, description, reports, etc.” (Ellis, 2003, p. 212). Cognitive classification, on the other hand, draws on the framework by Prabhu (1987, pp. 46-47), who classifies tasks into three groups: “information gap”, “opinion gap” and “reasoning gap” tasks. The final approach is psycholinguistic classification, which draws from the study by Pica, Kanagy and Falodun (1993) and deals with the tasks’ potential for language learning depending on interactional conditions (Ellis,

2003). This classification includes five types of tasks: “jigsaw”, “information gap”, “problem solving”, “decision making” and “opinion exchange” (p. 19).

Another most cited task typology is also proposed by Ellis (2003, 2012, 2020) based on a dual categorization of tasks. They include:

1. *focused* versus *unfocused* tasks
2. *real-world (real-life)* versus *pedagogic* tasks
3. *input-based* versus *output-based* tasks
4. *here-and-now* versus *there-and-then* tasks
5. *closed* versus *open* tasks

The task typology by Ellis (2003, 2012, 2020) examines whether a task requires employing receptive or productive skills (input-based vs. output-based tasks), has situational or interactional authenticity (real-world (real-life) vs. pedagogic tasks), orients learners to the use of certain linguistic features or language communicatively in general (focused vs. unfocused), provides learners with task related content while implementing task (here-and-now vs. there-and-then tasks) and has a single (or limited number of outcomes) outcome or not (closed vs. open tasks).

The typology by Ellis (2003, 2012, 2020) is employed in the present thesis as it provides a comprehensive analysis of tasks and makes pedagogical suggestions concerning how tasks should be sequenced based on their types across proficiency levels. What each task type involves is explained in the following sections.

Focused versus Unfocused Tasks

Ellis (2009, 2017, 2020) differentiates tasks as either *unfocused* (UN) or *focused* (FO) based on whether they orient task takers either to the use of certain linguistic features or language communicatively in general.

A focused task aims at fostering communicative language use (Ellis, 2003) and orienting learners to process predetermined, specific linguistic features (Ellis, 2009, 2012, 2020; Ellis et al., 2019) receptively or productively (Ellis, 2003). The linguistic features intended to be practised could be at morphological, syntactic, lexical, and pragmatic levels (Cobb, 2010).

Even if both focused tasks and exercises are designed to orient learners to particular linguistic features, focused tasks differ from exercises in that they still meet all the criteria for tasks, and the target linguistic feature aimed to be practised is not explicitly given or stated (Ellis, 2003, 2009, 2011, 2012, 2018; Ellis et al., 2019; Ellis & Shintani, 2014). Hence, focused tasks lead learners to employ the target forms incidentally (Ellis, 2003).

FO tasks include “structure-based production tasks”, “consciousness-raising tasks” (Ellis 2012, pp. 224-226) and “interpretation tasks” (Ellis, 1995, p. 88). FO tasks may cover specific language points to be purposefully examined, as in “consciousness-raising tasks” (Ellis, 2012, p. 226) or direct task takers’ attention to predetermined linguistic features in the spoken or written input to help them figure out the meanings of these features, as in “interpretation tasks” (Ellis, 1995, p. 88). They can also centre around general content such as “stories, pictures of objects, opinions about the kind of person you like” as in “structure-based production tasks” (Ellis, 2012, p. 226).

For the conceptualization of *structure-based production tasks*, Ellis (2003, 2012) and Ellis et al. (2019) draw on the study by Loschky and Bley-Vroman (1993). Loschky and Bley-Vroman (1993) determine three conditions for a structure-based production task that engages learners in using specific grammatical structures. A structure-based production task can 1) make the use of those structures “natural”

(i.e., target linguistic feature may occur naturally and repeatedly during task performance), (2) “useful” (i.e., target feature may be useful to accomplish the task) or (3) “essential” (i.e., the use of the target language point is a must to accomplish the task) (p. 132).

According to Ellis (2003), tasks can be designed to have a focus in two different ways. The first way is to design a focused task that can be completed only when a particular linguistic feature is used by the task takers (Ellis, 2003). Nevertheless, it can be challenging to design such tasks as there is always the possibility that task takers resort to communication strategies to avoid using the targeted feature (Ellis, 2003). The second way, on the other hand, is to make the targeted feature the content of the task, as in consciousness-raising tasks (Ellis, 2003). Consciousness-raising tasks make language points the content or their topic and lead learners to think and talk meaningfully about a specific language feature to successfully achieve the task goal (Ellis, 2003, 2012). Such tasks aim to engage learners in decoding language input to discover the rule behind it. As it has been observed that discovery-oriented activities have become commonly used in ELT textbooks (Ellis, 2016), consciousness-raising tasks can be adopted by the ESL/EFL textbooks.

The psycholinguistic rationale for employing FO tasks is explained through “skill-building theories”, “automatic processing”, “implicit learning”, and “focus-on-form” (Ellis, 2003, p. 144). FO tasks generate a context for pre-set linguistic points such as grammar structures or vocabulary items (Ellis, 2014) and draw the attention of the learners to the linguistic points they are not capable of producing correctly (Ellis et al., 2019), or somehow overlook or avoid using (Ellis, 2009). They provide means to teach grammar communicatively and support learners to produce

comprehensible and grammatically acceptable language (Nobuyoshi & Ellis, 1993). Another advantage of utilizing these tasks could be to check if learners have learned particular linguistic points taught (Ellis, 2003, 2015).

As opposed to *focused* tasks, *unfocused* tasks “are not designed in order to try to practice the use of any particular linguistic feature, grammatical structure.” (Ellis, 2020) but aim to enable learners to use language in general communicatively (Ellis, 2009, 2011, 2012, 2015, 2018; Ellis et al., 2019; Ellis & Shintani, 2014). The “natural” language to be utilized is broadly determined, and learners are not required to focus on or produce any particular linguistic feature (Nobuyoshi & Ellis, 1993, p. 204). That is, task designers or task-takers make no special effort to prioritise using any specific language point in the preparation or implementation phases for unfocused tasks (Nobuyoshi & Ellis, 1993).

Ellis (2009) recommends adopting a hybrid syllabus composed of focused and unfocused tasks. While Ellis (2003, 2009) and Long (1985, 2015) consider both focused and unfocused tasks, Willis (1996) suggests primarily unfocused tasks, and Skehan (1998) prefers only unfocused ones in a language syllabus.

Real-world (Real-life) versus Pedagogic Tasks

Another critical distinction concerns whether tasks have situational authenticity or not. Drawing on the classification of the communicative classroom tasks as “real-world” and “pedagogic” by Nunan (1989, p. 40), Ellis (2009) makes the same distinction between tasks based on whether they have interactional or situational authenticity. Real-world (RW) tasks correspond to the target tasks (Ellis et al., 2019) language users often perform in natural contexts outside school (Ellis, 2017) and can be called “real-life” (Ellis, 2020). These tasks have situational

authenticity (Ellis et al., 2019) as they imitate real-life contexts (Ellis, 2009), such as booking a hotel room and ordering a meal at a restaurant.

As opposed to RW tasks that target situational authenticity, pedagogic (PE) tasks seek interactional authenticity (Ellis, 2014, 2017, 2018; Ellis et al., 2019), which is achieved when the language produced during task implementation is similar to the one occurring in non-pedagogic conversation (Ellis, 2009).

Some examples of PE tasks are telling a story, using the images given, describing a picture while a partner is drawing it, listing the differences between two images, locating things on a map (Ellis, 2003), writing a news story based on the keywords and composing a text by ordering its sentences or paragraphs (Cobb & Lovick, 2007). Even if such activities are not likely to be formed in real life and, therefore, lack situational authenticity, they still have interactional authenticity as learners practice social actions such as “turn-taking”, “repairs of misunderstandings” (Ellis, 2017, p. 508) and “asking for and answering questions” (Ellis, 2003, p. 10).

It is noted that RW tasks are essential for language courses offered for specific purposes, whereas PE tasks may dominate the courses for general purposes (Ellis, 2018). PE tasks, which include interesting and familiar content, on the other hand, are beneficial for the courses for general purposes (Ellis et al., 2019).

Input-based versus Output-based Tasks

Input-based (IB) tasks include listening or reading tasks (Ellis, 2009, 2011, 2020; Ellis et al., 2019), which are also called “receptive tasks” (Ellis et al., 2019, p. 156) or “non-reciprocal tasks” (Ellis, 2012, p. 211). The accomplishment of such tasks depends solely on the comprehension of the input provided (Ellis et al., 2019).

Engaging learners in processing spoken or written information, IB tasks require task-takers to show comprehension in a non-verbal way “by an action,

selecting the right picture, finding the differences between two pictures, completing a map, or making a model” (Ellis, 2012, p. 211) or through minimal production at most (Ellis, 2016). IB tasks do not ask learners for language production, but they do not prohibit them from doing so (Ellis, 2009, 2012, 2013, 2014, 2015, 2017; Ellis & Shintani, 2014). Although IB tasks basically do not have a language production requirement, learners may engage in private or social speech (Ellis, 2009; Ellis & Shintani, 2014) using their L1 or L2 (Ellis et al., 2019). As only one party communicates information, these tasks are considered one-way tasks; yet learners may still negotiate for meaning if they feel the need (Ellis, 2009; Ellis & Shintani, 2014).

Within IB tasks, there should be a need to motivate task-takers to focus on the target items (e.g., lexical, grammatical) and to perform “form-meaning mapping” (Ellis et al., 2019, p. 49). As learners are required to utilize their (non) linguistic resources during task performance, they are expected to draw on their world knowledge (Ellis, 2009; Ellis & Shintani, 2014), L2 knowledge and contextual clues to decode the input and deal with the linguistic form (Ellis et al., 2019) through such strategies as “repetition, highlighting key items intonationally, utilizing contextual clues” (Ellis et al., 2019, p. 49). Teachers can also direct learners’ attention to form by giving feedback on their comprehension of the input (Ellis, 2009; Ellis & Shintani, 2014).

The use of IB tasks in textbooks can be advantageous for several reasons. It is claimed that such tasks support IL development by drawing learners’ attention to input in the target language while reading or listening (Ellis et al., 2019). In this way, learners can develop their abilities to comprehend input and grammatical resources to produce in the target language (Ellis & Shintani, 2014).

Prabhu (1987) and Ellis et al. (2019) argue that task-based learning should include IB tasks, especially for lower-level learners. According to Ellis (2012, 2017), input-based tasks provide comprehensible input that is necessary for L2 acquisition, as suggested by Krashen (1985), and help students learn L2 (second language / foreign language) in a similar way they acquire their L1 (first language) by being subject to input naturally (Ellis, 2014). Moreover, interactive IB tasks can be useful for young learners or beginners as these tasks facilitate incidental acquisition through negotiation of meaning (NoM) and the growth of implicit knowledge (Ellis, 2016). They may also be functional to familiarize learners with task-based teaching who might have difficulty adapting to a teaching model that uses language as a tool, not as an object (Ellis, 2017).

Unlike *input-based* tasks, *output-based* (OB) tasks engage learners in language production through speaking or writing (Ellis, 2009, 2011, 2012, 2015, 2017, 2018, 2020; Ellis et al., 2019). They may also involve minimal input (e.g., listening or reading), but output is always involved (Ellis, 2020). While IB tasks mainly involve one-way interaction, OB tasks generally seek to provide opportunities for two-way interaction (Ellis et al., 2019).

OB tasks are not advised to be integrated into the teaching of L2 to beginner-level learners since they are assumed to lack sufficient L2 knowledge to engage in free production (Erlam & Ellis, 2018). Hence it is strongly advised to begin with IB tasks with very low proficiency learners or beginner learners and delay OB tasks till they have acquired some language from practising IB tasks (Ellis, 2020).

Although a dual categorization (i.e., IB versus OB) is used for the use of language skills (i.e., listening, reading, speaking, writing), some tasks are integrative as they require learners to use two or more language skills (Ellis, 2009, 2011, 2012).

It has been observed that several researchers (Long, 1985, 2014; Robinson, 2015; Skehan, 1998; Willis, 1996) study primarily OB (production-based) tasks, thus paving the way for the conceptualization of task-based language teaching as concerning production-based tasks only (Shintani, 2012). However, Ellis (2003, 2012, 2020) focuses on both IB and OB tasks, providing a broader perspective of tasks.

Here-and-now versus There-and-then Tasks

Another classification of tasks is made based on the presence or absence of task input during task implementation. The contents of tasks are present in here-and-now (HN) tasks so that learners can revisit them whenever they want, whereas they do not have access to stimuli in there-and-then (TT) tasks (Ellis, 2012, 2020; Ellis et al., 2019). For instance, task-takers can keep the picture they are describing in a HN task (Ellis, 2006, 2012), but they talk about a film only after watching it without having access to it in a TT task (Ellis, 2012). A HN task can be made more difficult by making it a TT task by removing the task content needed to perform a task, such as a map and a list of the places required to be written on a map (Ellis, 2020).

Closed versus Open Tasks

Tasks can be classified based on the scope of the outcomes as open and closed tasks. Closed (CL) tasks are designed to have only one or a limited number of outcomes, namely solutions (Ellis, 2012, 2020; Ellis et al., 2019). Secondly, open (OP) tasks are the ones that have many possible outcomes, namely solutions (Ellis, 2011, 2020; Ellis et al., 2019). Typically, CL tasks correspond to “information-gap and reasoning-gap tasks”, whereas OP tasks to “opinion-gap tasks” (Ellis, 2012, p. 200).

OP tasks may be divergent or convergent depending on whether they require learners to reach a single outcome or give them the freedom to express different viewpoints (Ellis, 2003). While CL tasks may engage learners in such activities as decision making, open tasks may involve “making choices, surveys, debates, ranking activities and general discussion” (Ellis, 2003, p. 89). It is predicted that closed tasks lead to negotiation of meaning, while open tasks, converged versions, in particular, facilitate accuracy and complexity (Ellis, 2003).

Ellis (2020) recommends the use of particular types of tasks to be utilized for different proficiency groups (Table 3). The distribution of the task types can be seen in Table 3 below:

Table 3

Task Types Suggested for Different Proficiency Levels by Ellis

Type of Learner	Type of Task
Beginner	Pedagogic input based; closed; here-and-now
Intermediate	Pedagogic input based and output-based tasks; some focused tasks; mixture of closed/open and here-and-now and there-and-then tasks
Advanced	Mainly output-based tasks; focused tasks; real-life tasks; mainly open there-and-then tasks

Note. From *Using tasks in language teaching*, R. Ellis, 2020, September 8 [Webinar].

Cambridge Live Experience. Copyright 2020 by Cambridge Live Experience.

According to Ellis (2020), “pedagogic, input-based tasks; closed; here-and-now tasks” are more suitable for beginner-level language learners. When learners reach the intermediate level, using “pedagogic input-based and output-based tasks; some focused tasks; mixture of pedagogic and real-life tasks; mixture of closed/open and here-and-now and there and then tasks” would be the most effective for IL development (2020). As for the advanced levels, “mainly output-based tasks; focused

tasks; real-life tasks; mainly open there-and-then tasks” are considered the most suitable ones (2020).

As can be inferred from Table 3, Ellis (2020) categorizes proficiency levels into three groups only: beginner, intermediate and advanced levels. Even if task types for each proficiency group are not specified by Ellis (2020), it is indicated that some task types should precede the others as they are relatively easier to implement with lower-level proficiency groups. Figure 1 below demonstrates the order suggested to be followed as proficiency level increases.

Figure 1

The Sequencing of Task Types as Proficiency Level Increases



Figure 1 above shows that there must be followed an order from *pedagogic*, *input-based*, *here-and-now*, and *closed* tasks to *real-life*, *output-based*, *there-and-then*, *open* and *focused* tasks with the increase in the proficiency levels.

In the present study, the target student groups of the textbooks for Grade 9, 10, 11 and 12 are assumed to be elementary (A1+A2), pre-intermediate (A2+B1), intermediate (B1+/B2) and upper-intermediate (B2+...) proficiency levels, respectively. The task types within these textbooks will be examined based on the sequencing of task types as suggested by Ellis (2020).

Task Complexity and Sequencing of Tasks

Syllabus functions to coordinate decisions about instruction, testing, optimal stimulation of internal cognitive processes for IL development as well as the selection and organization of units of L2 instruction for any teaching context (Robinson, 2010, 2015). Syllabus design draws on the decisions on the units of classroom activities and the sequencing of the tasks to be implemented (Robinson, 2001a). Textbooks, designed based on the syllabus type adopted, can serve as an important means to logically order and effectively implement tasks when developed systematically and principally (Barrot, 2019).

Malicka (2020, p. 644) notes that determining “theoretically driven, pedagogically sound and researchable criteria for sequencing tasks” is critical for learner’s immediate performance during task implementation and the development of L2 in the long run. So far, there has not been a standard approach adopted to optimally sequence tasks for language instruction in the short or long run (Jackson & Suethanapornkul, 2013). Starting with the procedural syllabus put forward by Prabhu (1987) within the scope of *Bangalore Project*, several proposals have been made for task-based language syllabus (Robinson, 2015). Long (1985) suggests designing and sequencing pedagogic tasks based on the analysis of the needs of the learners in terms of the target tasks they need to perform in the real world. Another suggestion on how to sequence tasks was made based on the cognitive and performance demands tasks are predicted to make on the learners (Nunan, 1985). Nunan (1989) focuses on the ways to sequence tasks so that language learners improve their language skills to communicate or metacognitive skills to learn languages. Nunan (1985) recommends an order in which the activities start with comprehension-based activities, proceed to controlled production ones and end with interactive ones.

Another approach to syllabus design concerns task complexity. Task complexity is considered a critical design factor affecting performance and learning in the target language (Vasylets et al., 2017). The use of language activities necessitates the consideration of task complexity and sequencing them based on their degrees of complexity, as it is crucial to arrange task demands resulting from the task complexity dimensions at reasonable degrees for different proficiency levels (Samuda & Bygate, 2008). Various features of task components have been manipulated in several studies to see the effects of task design on learner's performance (Révész, 2011).

One way of classifying task features and sequencing tasks according to these features to facilitate language learning and increase language performance is to investigate task complexity from the cognitivist perspective. So far, the two leading perspectives of which assumptions have been accepted and studied concerning task complexity are *Limited Attentional Capacity* model by Skehan (1996, 1998, 2009) and *Cognition Hypothesis* (later developed into *Triadic Componential Model*) by Robinson (2001a, 2001b, 2005, 2007). Both perspectives assume that manipulating the features of tasks impinges on the attentional resources of the task takers, thus affecting their language production and IL development. Standing as the most widely cited proposals for the classification and the manipulation of language tasks, the proposals by these researchers will be examined in the following sections.

Limited Attention Capacity

Several studies were devoted to investigating task demands in terms of task difficulty (Brown et al., 2002; Norris et al., 1998; Prabhu, 1987; Skehan, 1996, 1998). Among them, Skehan's (1996, 1998, 2015) studies propose a more

comprehensive framework of task difficulty informed by the *Limited Attention Capacity* argument.

The main argument of LAC is that task performance is bound to attentional limitations affected by limited working memory size and that the pressure on limited resources leads to predictable consequences in L2 production (Wang & Skehan, 2014). The more difficult a task gets, the more significant the effect of attentional and memory limitations on L2 performance becomes (Wang & Skehan, 2014). Since attentional resources are limited, and attention is disproportionally allocated between three competing performance areas (i.e., accuracy, fluency and complexity) during the task implementation process, a trade-off occurs between these performance areas. Such a trade-off is more likely to be observed, especially in cognitively more demanding tasks that cause learners to push the limits of their attentional resources (Vasylets, 2017).

Task-based research is carried out to reveal how task characteristics and task conditions can minimize the influences of cognitive limitations (Wang & Skehan, 2014). LAC model claims that sequencing tasks based on a principled approach ensures a balance between accuracy and fluency in L2 performance by incorporating the previous restructuring into ongoing L2 use (Skehan, 1996).

Ideally, it is expected that language tasks are designed to have optimal conditions to enable learners to attend to the complex language structures and perform fluently without making errors (Skehan & Foster, 1999). However, according to LAC model, that attention is limited makes it unlikely to achieve fluency, accuracy and complexity simultaneously. In this case, the language performance during task implementation may hinge on some factors such as the

prioritizing decisions of task-takers, task features as well as task conditions (Skehan & Foster, 1999).

While the manipulations of some of the factors that influence L2 production during task implementation may inform us about some aspects of L2 performance at a time, Skehan and Foster (1999) note that the consequences of such manipulations should be explored for longer-term development as a balance needs to be sought among these performance areas so that L2 learners make progress in each of these areas. Therefore, one of the goals of LAC model is to facilitate a balanced development in three performance areas: complexity, accuracy and fluency (CAF).

Drawing on the LAC model, Skehan develops a *Task Difficulty Scheme*. Task difficulty relates to the amount of attention tasks require from the task-takers (Skehan & Foster, 2001). As content is prioritised over form, it tends to consume more from the attentional capacity when a task gets more difficult, thus leaving less amount of attention to be devoted to form. Less attention to form, in turn, results in decreased fluency, accuracy or complexity in L2 production (Skehan & Foster, 2001). As these performance indicators are susceptible to the cognitive demands of tasks, taking cognitive dimensions into consideration is crucial when designing tasks (Skehan & Foster, 2001).

Task Difficulty Scheme

The task-difficulty scheme is composed of three components: “formal factors (code complexity), content (cognitive complexity) and pressure to achieve communication (communicative stress)” (Skehan, 1996, p. 52). These areas are also called “language”, “cognition” and “performance conditions, respectively, and are claimed to affect attention allocation and task performance to some extent (Skehan &

Foster, 2001, pp. 195-196). The dimensions of the *Task Difficulty Scheme* are given in Table 4 (Skehan & Foster, 2001, pp. 194-195):

Table 4

Task Difficulty Scheme by Skehan and Foster

1. Code complexity	2. Cognitive complexity	3. Communicative stress
linguistic complexity and variety	cognitive familiarity	time pressure
vocabulary load and variety	familiarity of topic	scale
	familiarity of discourse genre	number of participants
	familiarity of task	length of texts used
	cognitive processing	modality
	information organization	stakes
	amount of ‘computation’	opportunity for control
	clarity of information given	
	sufficiency of information given	

Note. Adapted from “Cognition and Tasks”, by P. Skehan & P. Foster, in P.

Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 194-195), 2001,

Cambridge University Press. Copyright 2001 by Cambridge University Press.

Code Complexity. As given in Table 4, the first dimension in the *Task Difficulty Scheme* is “code complexity” (Skehan & Foster, 2001, p. 194). It relates to the linguistic demands tasks introduce (Skehan & Foster, 2001), specifically to syntactic and lexical difficulty and range (Skehan, 1996). Code complexity consists of two sub-categories – “linguistic complexity and variety” and “vocabulary load and variety” (Skehan & Foster, 2001, p. 194). If tasks necessitate “more advanced structures, or require the use of wider repertoires of structures, or greater densities of advanced structures”, they are considered more complex (Skehan & Foster, 2001, p. 195).

Cognitive Complexity. The second dimension in the scheme is cognitive complexity. “Cognitive complexity” concerns the task content and how it can be manipulated (Skehan & Foster, 2001, p. 195). It deals with the content of what is said

and relates to the conceptualization phase in the Levelt model (1989) (Skehan, 1996). It is further broken down into (1) “cognitive familiarity” and (2) “cognitive processing” (Skehan & Foster, 2001, p. 195).

Cognitive Familiarity. *Cognitive familiarity* concerns how much a task exploits ready-made or pre-packaged solutions (Skehan, 1996, 1998). It is also related to having access to the knowledge, which includes “relevant, already-organized material, and even solutions to comparable tasks” (Skehan, 1996, p. 52). This category consists of three factors: “familiarity of topic”, “familiarity of discourse genre” and “familiarity of task” (Skehan & Foster, 2001, p. 195).

Cognitive Processing. The second dimension of cognitive complexity is “cognitive processing”, which refers to the processing load during task implementation (Skehan, 1998, p. 270; Skehan & Foster, 2001, p. 195). Specifically, it denotes the amount of online computation needed for completing a task and to what extent a task-taker needs to actively scrutinize the task content (Skehan, 1996) to find solutions to the problems occurring during task implementation (Skehan, 1998). This dimension is further broken down into (1) “information organization”, (2) “amount of ‘computation’”, (3) “clarity of information given” and (4) “sufficiency of information given” (Skehan & Foster, 2001, p. 195).

Communicative Stress. The third main dimension in the *Task Difficulty Scheme* is “communicative stress” (Skehan & Foster, 2001, p. 195). It denotes the performance conditions under which tasks are implemented and the implications for processing (Skehan, 1998). It involves a set of elements that do not have a connection with code or meaning, but affect communication pressure (Skehan, 1996). These elements are (1) “time pressure”, (2) “scale” (i.e., number of

participants and length of texts used), (3) “modality”, (4) “stakes” and (5) “opportunity for control” (Skehan & Foster, 2001, p. 195).

The first element of the communicative stress dimension, “time pressure”, deals with the time limit within which a task is expected to be completed and the speed at which a task needs to be implemented (Skehan, 1996, p. 52). Task-takers may feel pressured to complete a task within a specified period, especially when the given time is short (Skehan & Foster, 2001).

The second potential source of stress is “scale”, which is concerned with “the number of participants” and the “length of texts used” (Skehan & Foster, 2001, p. 195).

The third element, “modality”, relates to the mode (i.e., listening, speaking, reading or writing), in which a task is instructed to be implemented (Skehan & Foster, 2001, p. 195). A task may require using only one skill or any combination of them, thus necessitating “real-time” or “slow-time” processing or both (p. 195).

“Stakes” relates to the orientation of learners to the tasks and is concerned with how important it is for learners to complete a task and perform it correctly (Skehan, 1996, p. 52). A task can be taken as a lightsome activity at the end of an exhausting class or serious work before an exam, affecting whether meaning or accuracy is prioritised (Skehan & Foster, 2001). The task stakes could be low when the task implementation process is given priority or if there are no consequences following it, while they can be high if accuracy is prioritized (Skehan, 1996).

Finally, “opportunity for control” deals with what extent task-takers can exert control over how a task is implemented through such opportunities as extra time to complete the task, an easy task content (Skehan & Foster, 2001, p. 195), asking clarification questions and negotiating meaning and the task itself (Skehan, 1996).

Cognitive Hypothesis

The main argument of the *Cognitive Hypothesis* is that tasks should be presented from simple to complex so that learners are supported to accomplish complex tasks in the L2 (Robinson, 2011). In this approach, the behaviour descriptions of L2 tasks are taken as the initial point in task design (Robinson, 2007) and task sequencing is performed through the design and implementation of certain task factors in their simple versions initially and then, gradually moving to more complex versions of these factors (Robinson, 2010).

Informed by “functional/cognitive linguistics”, “L1 developmental psychology”, and “SLA research” (Robinson, 2003, p. 47), CH makes claims on task complexity and sequencing, which can be grouped under five headings as (1) output, (2) uptake and interaction, (3) memory and retention, (4) automaticity and (5) individual differences as summarised below (Robinson, 2011):

1. *Output*: Higher cognitive demands posed in resource-directing dimensions decrease fluency, while increasing accuracy and complexity in L2 production during monologic task implementation. The same effects are expected during dialogic/interactive tasks, except for complexity, as interactants will be less likely to produce complex syntax.
2. *Uptake and Interaction*: Cognitively complex tasks increase interaction and negotiation of meaning.
3. *Memory and Retention*: As a result of increased attention to input during cognitively more complex tasks, the depth of processing will be relatively greater, and the input will be retained for a longer period of time.

4. *Automaticity*: Implementing tasks from simple to complex ones increases automaticity and induces effective organization of elements of complex task implementation in the target language.
5. *Individual Differences*: Complex tasks may potentially lead to more variation between learners in terms of language learning and task performance.

Triadic Componential Framework

It is claimed that if the cognitive demands of tasks are well-matched to the strengths and weaknesses of learners' abilities, then learners can successfully adapt to the demands of language tasks and make progress in language learning (Robinson, 2015). This goal requires an operational taxonomy for classifying the features of L2 tasks and the principles for sequencing the features in question (Robinson & Gilabert, 2007). TCF was created to respond to the need for the classification of the task features to be utilized when designing tasks and language syllabi targeting various learner profiles and learning settings (Robinson, 2011). Task complexity dimensions provide sequencing criteria to make appropriate prospective decisions regarding task units, whereas task condition and task difficulty factors inform online decisions regarding pair or group works (Robinson, 2001b).

Triadic Componential Framework (TCF), developed from CH, offers a detailed framework to study pedagogical tasks by scrutinizing several factors involved in the task process. "Triadic framework for investigating task influence on SLA" by Robinson (2001a) explicates the role of three dimensions in increasing the difficulty level of L2 tasks: *task complexity*, *task condition* and *task difficulty*. It is believed that successful L2 learning, and performance depend on the interplay between task demands introduced by task complexity, task conditions and the learner

capacity affecting how they perceive task difficulty (Robinson, 2003). Table 5 is devoted to three dimensions of TCF (Robinson, 2007, pp. 15-16).

Table 5

Criteria for Sequencing Pedagogic Tasks by Robinson

Task Complexity (Cognitive factors)	Task Condition (Interactive factors)	Task Difficulty (Learner factors)
(Classification criteria: cognitive demands) (Classification procedure: information theoretic analyses) (a) <i>Resource-directing variables</i> making cognitive/conceptual demands +/- here and now +/- few elements +/- spatial reasoning +/- causal reasoning +/- intentional reasoning +/- perspective-taking	(Classification criteria: interactional demands) (Classification procedure: behaviour - descriptive analyses) (a) <i>Participant variables</i> making interactional demands +/- open solution +/- one-way flow +/- convergent solution +/- few participants +/- few contributions needed +/- negotiation not needed	(Classification criteria: ability requirements) (Classification procedure: ability assessment analyses) (a) <i>Ability variables</i> and task-relevant resource differentials h/l working memory h/l reasoning h/l task-switching h/l aptitude h/l field independence h/l mind/intention-reading
(b) <i>Resource-dispersing variables</i> making performative/procedural demands +/- planning time +/- single task +/- task structure +/- few steps +/- independency of steps +/- prior knowledge	(b) <i>Participant variables</i> making interactant demands +/- same proficiency +/- same gender +/- familiar +/- shared content knowledge +/- equal status and role +/- shared cultural knowledge	(b) <i>Affective variables</i> and task-relevant state-trait differentials h/l openness to experience h/l control of emotion h/l task motivation h/l processing anxiety h/l willingness to communicate h/l self-efficacy

Note. Adapted from “Criteria for Grading and Sequencing Pedagogic Tasks”, by P.

Robinson, in M. D. P. Garcia-Mayo (Ed.), *Investigating Tasks in Formal*

Language Learning (pp. 7–27), 2007, Multilingual Matters. Copyright 2007 by

Multilingual Matters.

Three dimensions of the TCF given in Table 5, *task complexity*, *task condition* and *task difficulty*, are explicated in the following sections.

Task Complexity. The first dimension, ‘task complexity’, results from “the attentional, memory, reasoning, and other information processing demands imposed by the structure of the task on the language learner” (Robinson, 2001b, p. 29). Task complexity has to do with the intrinsic cognitive complexity of tasks (Robinson, 2008, 2015) and refers to the manipulable cognitive/conceptual demands they pose (Robinson, 2001a). It concerns intra-learner variation when implementing different versions of the same task (Robinson, 2001b, 2008, 2011; Robinson & Gilabert, 2007) as a consequence of the relatively lower or higher cognitive demands they make during task performance (Robinson, 2015).

Task complexity is examined in two dimensions: “resource directing” and “resource dispersing” (also known as “resource depleting”) depending on the cognitive/conceptual demands or performative/procedural demands made by the task variables (Robinson, 2001a, p. 294; Robinson, 2001b, p. 30; Robinson, 2007, p. 15). The presence or absence of these dimensions is indicated with + or – symbols, but they can also be thought of as a continuum, indicating the presence of less or more of a component (Robinson, 2001a, 2001b).

Task complexity is recommended to be the sole determiner of sequencing pedagogic tasks in syllabus design (Robinson, 2003, 2010, 2011). Only the task complexity dimensions are manipulable when designing tasks as it is possible to operationally map them from the intended real-world task to language task design (Robinson, 2010). It is noted that resource-directing and resource-dispersing elements of the *task complexity* dimensions may interplay and influence language output in measurable ways (Robinson, 2001b). Since it is not feasible to measure the *task condition* and *task difficulty* dimensions during task design stage, these are suggested to be excluded from a priori syllabus design (Robinson, 2010). However,

the task demands they pose are explicated within the framework, so that right in situ decisions can be made during instruction.

According to Robinson (2001a), designing and ordering tasks by increasing task complexity can positively affect language learning by activating cognitive resources, which progressively utilize learning mechanisms that induce the modification and restructuring of the IL. The influence of resource-directing and resource-dispersing factors within task complexity dimensions can differ, though. Increasing cognitive and conceptual demands through the manipulation of resource-directing dimensions is predicted to make task takers push output, whereas increasing task demands along resource-dispersing dimensions lead them to employ cognitive resources for attention and memory, to stimulate learning mechanisms and, finally, to incorporate input and modify output more than they would do during simpler tasks (Robinson, 2001a). Therefore, pushed and modified output is associated with the increasing task complexity in this theory. However, the synergetic effects to occur when factors from both resource-directing and resource-dispersing are manipulated should also be taken into account (Robinson & Gilabert, 2007). For instance, the assumed positive effect of increasing task complexity along one resource directing dimension can be minimized or questioned when combined with the increased complexity along a resource dispersing dimension. Hence, the interaction between these dimensions should be considered, and optimum resource allocation should be looked for to meet the linguistic demands posed by a task (Robinson, 2001b).

Resource-directing Factors. Resource directing factors concern the conceptual/communicative demands tasks pose (Robinson, 2015). They identify the task features based on the concepts required to be utilized or perceived during task

implementation, such as tense and location (Robinson, 2011; Robinson & Gilabert, 2007). The dimensions in this group refer to the developmental stages followed when acquiring language concepts in childhood (Robinson, 2010). They direct attention to ‘form-function/concept mappings’ and, therefore, foster IL development (Robinson, 2015, p. 9). Making tasks more demanding along resource-directing dimensions is predicted to lead task takers to utilize their cognitive resources and certain linguistic means in the L2 to accomplish a task (Robinson, 2001a, 2001b, 2007). Increased task complexity along resource-directing dimensions results in different learner performance in monologic and interactive tasks. Increased complexity positively affects accuracy and complexity, while negatively affecting fluency in monologic tasks. On the other hand, it decreases fluency but increases accuracy, comprehension checks and clarification requests in interactive tasks (Robinson, 2001a, 2003).

The resource-directing dimensions are “here and now”, “few elements”, “spatial reasoning”, “causal reasoning”, “intentional reasoning”, and “perspective-taking” (Robinson, 2007, pp. 15-16).

The first resource-directing dimension is *here-and-now*. It concerns whether tasks necessitate referring to “events happening now in a shared context (here-and-now) to orient learner attention to morphology for conveying tense and aspect in the present” or “events happening elsewhere in time and space (there-and-then)” (Robinson, 2010, p. 247; Robinson, 2015, pp. 15-16). +HN condition involves the narration of events occurring in a shared context now (Robinson, 2011). In the *there-and-then* condition (-HN), on the other hand, the event learners are required to narrate is visually unavailable in a shared context (Robinson, 2010, 2015). The latter condition is considered cognitively and conceptually more demanding as more effort

is made for conceptualization and greater demands are imposed on memory (Robinson, 2010, 2015).

In previous studies, -HN condition was called as “there-and-then” (+TT) condition (Gilabert, 2007, p.223; Robinson, 1995, p.140), - “contextual support” (Choong, 2011, p.11; Révész, 2007, p.14) or *-immediacy* (Iwashita et al., 2001, p.414). Besides, -HN condition was operationalized in different ways. While some studies (Gilabert, 2007; Gilabert et al., 2009; Ishikawa, 2006; Robinson, 1995, Robinson et al., 1996) examined the effect of both the presence or absence of a non-verbal material (e.g., a cartoon/comic strip, picture, photo, video) as well as tense use (present versus past) to examine HN dimension, some focused only on the presence or absence of a non-verbal material (Choong, 2011; Iwashita et al., 2001; Révész, 2007; Révész et al., 2011; Skehan & Foster, 1999) that would function to provide a shared context. The factors that generated HN condition were also differently operationalized as allowing task-takers to see one another or their avatars during online interaction in one study (York, 2019) and as the presence or absence of a reader in another (Ishikawa, 2006).

The second dimension, *few elements*, involves differentiating between and identifying one or more elements/objects among many others (Robinson, 2007). This dimension concerns if there are few, easily distinguishable elements (+FE) as opposed to several similar ones to be dealt with in a task (-FE) (Robinson, 2007, 2008; Robinson & Gilabert, 2007).

Previous research that investigated the FE dimension operationalized it as the number of objects (York, 2019), characters (Donate, 2018; Kang & Lee, 2019; Sasayama, 2011; York, 2019), location and location characteristics (Lee, 2018), criteria for decision making (Kuiken & Vedder, 2007), landmarks (Gilabert, 2007),

landmarks with the axis (or exes) they move along (Gilabert, 2007), landmarks and stops on a map (Solon et al., 2017), projects to allocate money to (Révész, 2011) as well as hotel and client characteristics (Malicka, 2014a).

Another resource-directing dimension is *spatial reasoning* (SR). Tasks that involve complex spatial reasoning engage learners in activities such as "describing how to move, and in what manner, from point A to point E, by way of intermediary landmark points B, C and D" induce the employment of the linguistic forms to describe motion events (Robinson, 2011, p. 16). These tasks are predicted to increase learners' awareness of lexicalization patterns in L2 to describe the motion events (Robinson, 2011). Easily identified and mutually known landmarks can be provided for spatial location in a simple task, but this support is absent in the complex version (Robinson, 2007, 2008; Robinson & Gilabert, 2007). In the previous studies with +SR conditions, activities such as seating people at a table for different events (Nuevo, 2006) and decorating a room when the task-takers' views of the scene were not fixed (York, 2019) were used. It must be noted that task-takers are expected to think about the spatial positions of people, objects etc., to assume that a task requires spatial reasoning instead of simply describing their positions when space references are already provided (Levkina & Gilabert, 2014).

Causal reasoning (CR) dimension concerns whether a task requires "simple information transmission, versus reasoning about causal events and relationships between them" (Robinson, 2007, p. 18; Robinson, 2008, p. 7; Robinson & Gilabert, 2007, p. 165). Some of the previous studies with +CR condition engaged participants in ordering pictures to create a plot (Choong, 2011; Kim, 2009; Nuevo, 2006), selecting the right description of the causal events manipulated in terms of directness, intentionality and agency (Choong, 2014), reporting a boss about the hypothetical

changes in human relationships among staff based on job mistakes at a workplace (Ishikawa, 2008), making judgments about whether to hold a specific group of people responsible for an event (Hsu, 2020), apologizing to hotel clients providing explanations and reasons for a state of affairs (Malicka, 2020) as well as making and justifying a series of decisions for the actions to take and the sequence of these actions.

Intentional reasoning (IR) dimension examines if a task requires reasoning about the intentions of people in performing an action (+IR) or not (-IR). Tasks with intentional reasoning cause learners to reason about the intentions, beliefs and desires of people as well as the relationships between these (Robinson, 2007, 2008; Robinson & Gilabert, 2007). These tasks are predicted to enable learners to notice and internalize specific linguistic features to accomplish tasks (Robinson, 2015). Among these linguistic features are the psychological and cognitive state terms such as *think* and *believe* that are used to refer to other minds as well as the complement constructions accompanying their use such as “she suspected, wonders” (Robinson, 2011, p. 16). So, the tasks that require intentional reasoning necessitate more than a simple description of people’s actions (Baralt, 2010).

In the studies with +IR conditions, tasks required understanding and predicting the motives and intentions of the characters (Robinson, 2000), choosing the best room for hotel clients (Malicka, 2020) as well as seating people at a table for different events taking their intentions into account (Nuevo, 2006). In addition, drawing on Robinson’s conceptualization of IR and borrowing from cognitive psychology, Awwad (2017), redefined the term as “reading other people’s thoughts and understanding their desires, beliefs and motives to draw true conclusions about what they intend to do, why and what consequences follow” (p. 44).

Finally, *perspective-taking* (PT) dimension concerns if a task causes learners to adopt only one first-person perspective on an event (-PT) or assume several second- and/or third-person perspectives (Robinson, 2007, 2008; Robinson & Gilabert, 2007) by using different pronominal references. One study that focused on PT engaged learners in telling a story as if it happened to them or from somebody else's point of view (Iwashita et al., 2001).

Resource-dispersing Factors. Resource-dispersing dimensions concern the performative and procedural demands tasks make on cognition (Robinson, 2007, 2010, 2011, 2015). They include “planning time”, “single task”, “task structure”, “few steps”, “independency of steps” and “prior knowledge” (Robinson, 2007, pp. 15-16).

Increased complexity on resource-dispersing dimensions is predicted to cause greater demands on attention and working memory (Robinson, 2001b) and, therefore, to decrease the accuracy and complexity of linguistic production (Robinson, 2008). Increasing complexity along these dimensions does not lead L2 users to direct their resources to use specific linguistic features as in the resource-directing dimensions, but diminishes the memory and attentional resources and, therefore, negatively affects the language performance in terms of CAF areas (Robinson, 2001b) and develops language users' skills to implement the task by triggering the processing conditions these language users can be subject to in daily life (Baralt et al., 2014).

Although increasing task complexity has an inverse relationship with complexity in language production, it is predicted to lead to more negotiation and interaction (Robinson, 2001b). It may also disperse the attention of the task-taker over several (non)linguistic aspects of a task and enable greater control over and automatic access to present IL linguistic resources (Robinson, 2015). Considering the

assumptions of the CH for resource-dispersing dimensions, when language users have, for instance, less background knowledge (-prior knowledge) relevant to the tasks they deal with, it can be expected that this condition will decrease their attentional resources to fulfil the task demands, thus causing poorer language performance (Robinson, 2001b). Still, it will initiate more negotiation and interaction, thus increasing their opportunities to practice L2 and develop IL.

The first resource-dispersing dimension is *planning time* (PT). It has to do with whether learners are provided with planning time or not (Robinson, 2007, 2008; Robinson & Gilabert, 2007). Giving planning time increases resource availability (Robinson, 2007, 2008), while removing it disperses task-taker's attention over task aspects to be dealt with simultaneously (Robinson & Gilabert, 2019).

Some common task planning conditions include *pre-task planning* (also named strategic planning), *during-task planning* (also called online planning), *task repetition* (also named rehearsal) and *no planning*. Planning time was operationalized in a variety of ways in the previous research. Iwashita et al. (2001), for instance, adopted a binary system of +/-PT before the implementation of tasks. In some studies, planning was manipulated as pre-task, online, and no planning conditions (Ellis & Yuan, 2004; Yuan, 2001; Yuan & Ellis, 2003). While no time limit was determined for online planning, the durations of pre-task planning varied in these studies.

In another study (Khatib & Farahanynia, 2020), PT conditions were varied as task repetition (performing the task twice without pre-task planning), strategic planning (10 minutes given each time for pre-task planning and task implementation), joint planning (repeating the same task after 10-minutes of pre-task planning) and no planning (no task repetition and no pre-task planning).

The study by Tabari (2018) determines three task conditions as pre-task planning, online planning as well as pre-task and online planning. Finally, four planning conditions were determined by Ahmadian and Tavakoli (2010, p. 35) as “careful online planning without task repetition”, “pressurized online planning with task repetition”, “careful online planning with task repetition”, and “pressurized online planning, no task repetition”. In the study, task repetition concerned whether the same task was implemented one week later or not, while online planning was operationalized as pressurized (with time restriction of 6 minutes) or careful (without time restrictions).

The second resource-dispersing dimension is the *single task* (ST). This dimension concerns if a task necessitates just one thing to complete or two or more things to be dealt with simultaneously (Robinson, 2007, 2008, Robinson & Gilabert, 2007). For instance, if the task is answering the phone in +ST condition, answering the phone can be accompanied by checking the weather simultaneously in -ST condition (Robinson & Gilabert, 2019). Dual tasks (-ST) are predicted to result in decreased accuracy, fluency, and complexity in language production (Robinson & Gilabert, 2019).

Task demands along ST dimension can be increased by engaging task-takers in more than one action during task implementation such as watching and retelling the story in the video (Skehan & Foster, 1999), taking comprehension and vocabulary tasks simultaneously (Peters, 2007), determining the route on a map and simultaneously giving directions for the destination (Robinson et al., 1995), describing a photo without the photo being present (Révész, 2009) and reacting to colour change while implementing another task (Révész et al., 2015).

Another resource-dispersing dimension is the *task structure* (TS). It is related to the presence or absence of a clear structure provided to support learners to determine what steps are involved in completing the task (Robinson, 2007, 2008; Robinson & Gilabert, 2007). In previous research, task structure was operationalized as whether or not learners were provided with further guidance on how to implement the task with a worksheet to organize the interactive discourse (Alwi et al., 2012).

While the dimension, *few steps* (FS), concerns whether one or few steps are involved in a task or many steps to complete (Robinson, 2007, 2008; Robinson & Gilabert, 2007), the dimension of *independency of steps* (IS) specifies if a task consists of activities to be completed in a strictly chained sequence following one after the other or not (Robinson, 2007, 2008, Robinson & Gilabert, 2007).

Finally, *prior knowledge* (PK) is related to the background information provided to task-takers (Robinson, 2007, 2008, Robinson & Gilabert, 2007). +PK condition was operationalized in the previous research as *familiar area* (Robinson, 2001b), *similar topic* (Robinson et al., 1995), *task familiarity* (Arslanyılmaz & Pedersen, 2010), *knowledge of a domain* (Adams & Alwi, 2014), *content familiarity* (Hardy & Moore, 2004; Urwin, 1999) and *form* (e.g., organization of lectures) familiarity (Urwin, 1999) depending on the context of the studies in the literature.

Task Conditions. Task conditions involve “two categories of task characteristics that affect the nature and amount of interaction between participants in performing real-world tasks and pedagogic task versions of these” (Robinson, 2015) called participation and participant dimensions (Robinson & Gilabert, 2007) that make interactional and interactant demands, respectively (Robinson, 2015).

Participation variables that make interactional demands consist of “open solution”, “one-way flow”, “convergent solution”, “few participants”, “few

contributions needed” and “negotiation not needed” (Robinson, 2007, pp. 15-16).

Secondly, participant variables that make interactant demands include “same proficiency”, “same gender”, “familiar”, “shared content knowledge”, “equal status and role” and “shared cultural knowledge” (Robinson, 2007, pp. 15-16).

Task conditions concern the situational setting and the conditions in which tasks are implemented (Robinson, 2008) and influence how and to what degree task-takers interact (Robinson, 2010). Task conditions are not recommended to be taken as a basis for sequencing tasks (Robinson, 2007). Instead, they are suggested to be determined a priori based on the needs analysis conducted at the start of the instructional design process and replicated every time a more cognitively complex version of a task is implemented so that language practice will be able to be transferred to a real-world context eventually (Robinson, 2010).

Task Difficulty. Task difficulty constitutes the third dimension in the TCF. It refers to learner-related factors and deals with affective and ability variables (Robinson, 2001a, 2001b, 2015; Robinson & Gilabert, 2007), which add to the perceived difficulty of the tasks (Robinson, 2008). Specifically, it has to do with how learners perceive the task demands and is affected by both the task complexity and task conditions (Robinson, 2003).

The cognitive abilities of the learners and other individual differences may affect learner performance during task implementation (Robinson, 2015). Task difficulty can become more prominent when the complexity level of a task is higher; therefore, it can account for learner variation in task performance (Robinson, 2011).

Individual differences among learners are predicted to affect the perception of *task difficulty* and interact with *task condition* and *task complexity* dimensions (Robinson, 2011). When combined with certain task conditions, learner factors can

predict to what degree the engagement in task participation is successful (Robinson, 2015). Although no fixed relationship between *task complexity* and *task difficulty* can be expected as learners may show variety in such aspects as intelligence, aptitude, or motivation, it is still maintained that there might be a connection between *task complexity* and learners' perceptions of *task difficulty* since task difficulty variables are claimed to affect the learners' perceptions of tasks (Robinson, 2001a, 2001b). Acknowledging that what learner factors involve is still unclear, Robinson notes that some task difficulty dimensions have not been sufficiently studied yet for their effects on task performance (Robinson, 2015). Research into the impact of learner factors to satisfy the cognitive and interactional demands tasks pose would help increase the likelihood of L2 development and make more appropriate decisions on differentiating instruction (Robinson, 2007).

The temporality of learner factors and the existence of a plethora of potential factors to affect them makes it harder to predict the learner factors when designing language materials and syllabi (Robinson, 2001a, 2001b). Therefore, these factors are not considered for syllabus design (Robinson, 2015). However, taking learner factors into account is crucial for online methodological decisions while implementing language syllabi (Robinson, 2001a, 2001b).

SSARC

Robinson (2015, p. 12) suggests two task sequencing principles underpinning the SSARC (“stabilize”, “simplify”, “automatize”, “restructure”, “complexify”) model developed out of CH. The first principle states that only the cognitive task demands should be taken as a basis for task sequencing. The second specifies that resource-dispersing dimensions should be made complex first, and then resource-directing ones.

Task complexity involves several options for manipulating the task demands to start with the simpler tasks and gradually proceed to the most complex version (Robinson, 2001a). The first step of SSARC is to design a simple task on both resource-directing and resource-dispersing dimensions so that performance employs “stable (SS) ‘*attractor state*’” of the present IL (Robinson, 2015, p. 10). Second, task is made complex along resource-dispersing dimensions to enable faster access to “automatization (A)” of the IL (p. 11). Third, task is made more complex along both dimensions to enable “restructuring (R)” of the IL, develop new form-function/concept mappings and present maximum “complexity (C)” to destabilize the current IL system (p. 11). Robinson (2015, p. 101) shows how SSARC formulated above can be actualized with an example (Figure 2).

Figure 2

A Sample Task Increased in Complexity Based on the Sequencing Principles of SSARC Model

Dimensions of complexity	Simple 1	2	3	4	Complex 5
planning time (before speaking)	+	–	–	–	–
single task (route marked)	+	+	–	–	–
prior knowledge (a familiar area)	+	+	+	–	–
few elements (a small area)	+	+	+	+	–

(simplified data/map) (authentic data/map)

Note. From “The Cognition Hypothesis, Second Language Task Demands and the SSARC Model of Pedagogic Task Sequencing”, by P. Robinson, in M. Bygate (Ed.), *Domains and Directions in the Development of TBLT: Plenaries from a Decade of the International Conference* (pp 87-133), 2015, John Benjamins. Copyright 2015 by John Benjamins.

In Figure 2, a direction-giving map task is demonstrated, which is increased in complexity following the sequencing principles of SSARC. The real-world target task is to give directions to a driver by using a map so that they find their way in an unfamiliar region. The simplest task, corresponding to SS stage (i.e., simple/stabilizing IL), is designed to be simple on each task complexity dimension. Task taker is given planning time to give directions on a map displaying a small area with a marked route as well as few roads and buildings/landmarks. Gradually, the task is made more complex on resource-dispersing factors to automatize access to IL (A) by excluding planning time (-PT) on the II. version and excluding both planning time (-PT) and route marking (-ST) on the III. version. The map is changed to be unfamiliar to the task takers on the IV. version (-PK), thus increasing the performative demands. Finally, when the complexity of the task is also increased along the resource-directing dimension (-FE), it approximates the real-world target task with an authentic map displaying an unfamiliar region and more elements (e.g., roads, landmarks). The fifth and the last version (V. version) engages learners with more complex language and leads to “restructuring and complexifying (RC)” of IL (Robinson, 2015, pp. 18-19).

Criticism against Triadic Componential Framework

Triadic Componential Framework (TCF) is widely used in studies investigating task complexity, conditions and difficulty since it proposes a comprehensive framework for task demands. Nevertheless, the dimensions of TCF and the predictions based on these dimensions have drawn a lot of criticism that needs to be considered in task-based research drawing on this framework.

Lee (2018) notes that one of the main issues that can be considered as limitations of the proposal by Robinson is the absence of standard operationalization

of complexity factors. Many researchers reported an unsystematic approach to operationalising the concepts related to task complexity in the existing literature (Awwad, 2017; Awwad et al., 2017; Jackson & Suethanapornkul, 2013; Lee, 2018; Malicka, 2014a; Nuevo, 2006; Sasayama, 2016). It is also maintained that the presence of several variables leads to too many interpretations of these variables (Malicka, 2020).

For Skehan (2016), two resource-dispersing variables in Robinson's framework, *planning time* and *task structure*, can be taken as resource-directing variables. Besides, the dimensions, *number of elements* and *reasoning*, have some overlaps, which makes it difficult to distinguish them from one another (Lee, 2018; Skehan, 2016). Moreover, Long (2015) argues that some task complexity dimensions (e.g., *planning time* and *prior knowledge*) may function as task conditions, and objective measures of some task complexity factors may be elusive.

Finding the conceptualization of the *planning time* dimension problematic, Wang and Skehan (2014, p. 159) maintain that planning involves several processes and goals such as “planning-for-task-interpretation”, “planning-for-organisation”, and “planning-as-rehearsal” that would affect performance aspects differently and make the construct impossible to be treated as monolithic. Moreover, for both strategic and online planning, the conclusions regarding their effects on L2 production can be reached only through inference following the task implementation without any direct evidence as to whether planning is the factor that has changed task performance or even whether planning took place or not (Skehan & Foster, 2005).

Another problem with the operationalization of task factors in Robinson's framework concerns *few elements*. Malicka (2014a) mentions the lack of clear guidelines to determine the sufficient number of characters, considerations or

elements to label a task simple or complex. The dimension of FE is generally taken intuitively, assuming that the increase in the number of elements makes a task more complex (Malicka, 2014a; Michel, 2011). However, previous research did not find homogenous results as to the effect of the addition or reduction of the elements in task content (Malicka, 2014a; Michel, 2011), which might be due to the differences in the operationalization of the FE dimension.

Malicka (2014a) states that most of the previous studies designed tasks in a way that one has two elements (+FE) while the other has four or six elements (-FE). However, this distinction was made arbitrarily instead of drawing on empirical evidence or objective principles (Malicka, 2014a). It is not clearly specified what number of elements makes a task have +FE or -FE condition (Lee, 2018). Samuda and Bygate (2008) note that it is only possible to interpret the effect of the few elements dimension relatively with reference to the quantity of the elements in a specific task. Besides, their effect can differ based on the degree of their similarity/distinctiveness as well as their capacity to create a related group of elements.

Another issue to be considered regarding the FE dimension is that it is difficult to make sure how many of the given elements are salient to the task taker to deem a task simple or complex (Lee, 2018). It may be another reason why previous research did not find homogenous results as to the effect of adding or reducing elements in task content (Malicka, 2014a). For instance, the study by Sasayama (2016) demonstrates that when the differences between the versions of the same task which are manipulated along the FE dimension are greater, their effects are noticeable.

Révész (2007, 2009) and Malicka (2014b) notes that although the HN dimension and the contextual support condition are treated as the same task complexity factor in some studies, these factors seem to differ in Robinson's task framework (2001a, 2005). Révész (2007, 2009) notes that conceptualizing HN as (non)existence of contextual support seems to be problematic. She argues that the role of contextual support in task performance is disputable due to lack of clear explanations in the previous research, maintaining that it is not possible to distinguish whether it is the differential linguistic difficulty (the use of past tense versus present tense) or the (non)existence of contextual support that affects the task performance as long as these two factors are not separately examined.

Another issue concerning contextual support condition is that when it is available, there is a single demand of describing and linking the series of events to narrate; however, its absence may necessitate both remembering the details of the story (procedural/performative demands) and narrating it in the past tense (cognitive/conceptual demands). However, procedural/performative demands, and cognitive/conceptual demands are associated with different categories of task complexity dimensions (resource dispersing versus resource directing). As the absence of contextual support might increase procedural/performative demands, Révész (2007, 2009) concludes that contextual support is more likely to be a resource-dispersing dimension, rather than a resource-directing one. Likewise, Malicka (2014b) maintains that HN should be considered resource-directing dimension, whereas contextual support resource-dispersing dimension.

Due to similar concerns, Michel (2011) that the interpretations of the learner performance based on the manipulation of HN condition is questionable. Hence Révész (2007) suggests that future research studies the effect of contextual support

independently from linguistic condition, which is related to resource directing dimension. If the factors are not controlled well, this situation may jeopardise the validity and reliability of a study.

Skehan (2016, p. 40) finds the so-called difference between resource-directing and resource-dispersing variables unclear and slightly arbitrary. Also, Samuda and Bygate (2008) maintain that the resource-dispersing (i.e., resource-depleting) dimensions seem to better fit into the task conditions category rather than task complexity.

Another point worthy of consideration is that task complexity has generally been operationalized by designing relatively simple and complex versions of the same task along a task complexity dimension, which is predicted to affect the cognitive load during task implementation (Révész et al., 2015). Nevertheless, it may not be feasible to compare the degree of cognitive complexity a factor brings about to a task (Sasayama, 2015). For instance, even if a researcher designs a task with its simple and complex versions, they may still not differ in their cognitive load enough to lead to variation in language performance (Sasayama, 2015, 2016).

It is debatable that to what extent each dimension of task complexity causes higher task complexity (Wang & Skehan, 2014). Furthermore, the cognitive load of each cognitive complexity dimension may not be the same (Sasayama, 2015). However, it is not clearly understood from Robinson's framework whether all task complexity dimensions are equally important or not (Samuda & Bygate, 2008).

Robinson (2007) acknowledges that the relationships between task characteristics and their combined effects in terms of task complexity remain uncertain and need to be studied to theoretically and empirically inform the design of language tasks and programs. Sasayama argues that although the concept "cognitive

task complexity” has drawn heavy attention in the field of TBLT, there has not been a consensus on its impact on L2 production, interaction or learning yet (2016, pp. 231-232). Although the predictions of the CH by Robinson as to task complexity have been widely studied, the findings are far from providing satisfactory evidence that supports or rejects them (Malicka, 2014a). It has been observed that Robinson’s predictions were tested in several studies, which mostly measured oral performance and revealed inconsistent results (Awwad, 2017; Frear, 2013; Jackson & Suethanapornkul, 2013; Malicka, 2014a; Révész, 2007; Vasylets & Gilabert, 2013; Wang & Skehan, 2014). As for monologic tasks, only the predictions for manipulating resource directing factors are valid, while the manipulations of both resource-directing and resource-dispersing dimensions apply when it comes to dialogic tasks (Gilabert, 2009)

Acknowledging that although the inconsistency between the findings may be attributed to several factors, Sasayama (2016, p. 232) accepts that the main concern pertains to the “actual versus designed” influence of the independent variable, that is, task complexity, as also noted by Robinson (2001a). Therefore, task complexity dimension needs to be clarified as the predictions regarding the manipulations of task complexity may not actually change the complexity levels (Skehan, 2016). Besides, Robinson notes that more evidence is needed for the short- and long-term effects of various task sequencing models on the acquisition and performance of L2 (2011).

It must also be acknowledged that the manipulation of the factors listed in the TCF by Robinson might not be the only reason behind task performance. An important problematic issue relates to how tasks are interpreted by task-takers (Lee, 2018; Long, 2015; Skehan, 2016). The way task-takers interpret and perform tasks can be different from the predictions due to individual factors. In addition to

individual factors, task conditions, which may also change the cognitive load of a task to a great extent (Lee, 2018; Sasayama, 2015), need to be considered as much as task complexity. Moreover, L2 communication tasks may be influenced by several other factors that make it more difficult to detect the underlying reason(s) behind learner performance (Sasayama, 2015). Moreover, how different dimensions within TCF interact with one another is not specified sufficiently (Lee, 2018).

To address this problem, several researchers attempted to diversify the evidence gathered during and after task performance by using different measures such as self-ratings of learners for task difficulty (Baralt, 2010; Donate, 2018; Révész et al., 2015; Sasayama, 2016; Yoon, 2017), subjective time estimation (Baralt, 2010; Sasayama, 2016), dual-task methodology (Révész et al., 2015; Sasayama, 2016) and expert judgement (Révész et al., 2015).

It is worthy of consideration that not only the operationalizations of simple and complex versions of the same task but also the absence of systematic selection and definition of outcome measures can be problematic (Lee, 2018). Moreover, the interpretation of the results of the task manipulations may differ (Pallotti, 2019). For instance, Pallotti (2019) questions how the findings that do not support the predictions of CH should be interpreted since they may be taken as a piece of evidence against CH or interpreted as that the manipulated factor does not increase task complexity.

Although TCF has drawn a lot of criticism so far, it is still valuable in that it brings up important questions to explore and improve language tasks (Samuda & Bygate, 2008). CH, from which the TCF was developed, focuses explicitly on cognitive demands and separates controllable and manipulable demands from non-controllable and non-manipulable ones (Malicka, 2020) and is considered to be the

most effective deductive approach for task demands (Skehan, 2016). Moreover, drawing on CH, the SSARC by Robinson is regarded as a conceptually more developed task sequencing model than existing proposals (Malicka, 2020) and the only fully fledged model present for task sequencing (Malicka, 2014a).

Considering the strengths and weaknesses of the TCF by Robinson, which have been mentioned within the present thesis, it would be better to take Robinson's framework as a preliminary hypothesis to be modified and refined as some researchers suggest (Samuda & Bygate, 2008). Therefore, task variables in TCF as well as the principles of SSARC would be practical as a base to examine task complexity and sequencing. Some dimensions would need redefining considering the previous criticisms and the scopes of the studies to be conducted, though.

Comparison of LAC and CH

Cognitively oriented task-based research mostly drew on the theoretical approaches to task demands by Skehan and Robinson (Adams & Alwi, 2014; Arslanyılmaz & Pedersen, 2010; Awwad, 2017; Awwad et al., 2017; Cho, 2011; Choong, 2014; Donate, 2018; Gilabert, 2007; Gilabert et al., 2009; Hardy & Moore, 2004; Hsu, 2012; Larson, 2020; Lee, 2018; Luo, 2007; Moradi & Farvardin, 2019; Révész, 2011; Riccardi, 2014; Sasayama, 2015; Tabari, 2018; Vo, 2018; Yıldız, 2017).

LAC and CH take tasks as means to facilitate L2 learning and propose suggestions to analyse task demands to design tasks in a principled way. Besides, even if neither Robinson nor Skehan openly acknowledges that their theoretical frameworks address only oral production, it is claimed that the frameworks by both Robinson and Skehan were initially developed to target speaking performance (Vasylets, 2017). However, they adopt different views regarding the kind of

language use and language learning opportunities that occur while implementing tasks in L2, as well as how L2 learners allocate their attentional resources during the task implementation process (Vasylets, 2017).

One difference between these frameworks is that while CH considers task complexity accountable for increased accuracy and complexity, LAC takes both task characteristics and task conditions into consideration to account for L2 performance (Wang & Skehan, 2014) and suggests that priorities in performance conflict when attentional resources are limited (Skehan & Foster, 2012). Another difference lends itself to how attentional capacity is taken. Skehan (1996, 1998) focuses on the limitations in the L2 learners' attentional capacity and claims that increased task complexity along each dimension leads to decreased performance in complexity, accuracy and fluency (CAF) at the same time, whereas Robinson (2001b) adopts a multiple-resource model of attention and emphasizes attentional capacity instead, which affects each performance area differently. Besides, one of the most significant differences between two approaches would be that LAC does not make suggestions specifically for syllabus design while CH proposes SSARC to sequence and grade tasks to promote IL development.

Robinson (2015) lists some of the differences between CH and LAC (also known as the *Trade-off Hypothesis*):

1. The *Trade-off Hypothesis* by Skehan does not seek to propose a model to sequence pedagogic tasks, design language syllabus, or conduct empirical research to test its predictions.
2. The *Trade-off Hypothesis* asserts that learners can pay attention to either accuracy or complexity separately, not simultaneously. However, CH

claims complex tasks are likely to increase both accuracy and complexity on monologic tasks.

3. The *Trade-off Hypothesis* does not make predictions about how the amount and nature of interaction get affected by the task demands as CH does.
4. CH distinguishes task complexity factors and makes predictions for the task demands they each may create. Besides, it argues that both general and specific measures of accuracy and complexity should also be employed to test the predictions of the hypothesis for simple and complex tasks.
5. CH distinguishes between *task complexity*, *task conditions* and *task difficulty*.

Although Skehan and Robinson adopt opposing positions regarding the effects of increased cognitive load on learners' performance during L2 task implementation, some of the components of the *Task Difficulty Scheme* developed out of LAC and TCF developed out of CH refer to similar issues.

The *cognitive familiarity* category in *Task Difficulty Scheme* (Skehan, 1996, 1998; Skehan & Foster, 2001) deals with to what extent a task contains ready-made or pre-packaged solutions and consists of three factors as *topic familiarity*, *familiarity of discourse genre* and *familiarity of task*. Likewise, the *prior knowledge* dimension by Robinson is defined in a way to refer to background information provided to learners in general (Robinson, 2007, 2008; Robinson & Gilabert, 2007). It could be concluded that the concept of PK may also refer to three categories of cognitive familiarity as defined by Skehan (1996, 1998) as it can cover a broad range of concepts related to familiarity.

Scale in Task Difficulty Scheme is concerned with ‘length of texts used’ and ‘the number of participants’ implementing a task (Skehan & Foster, 2001). The second subcategory of scale category can be claimed to have overlap with some of the *few participants* variable of the task conditions dimension of TCF by Robinson (2001b, 2003, 2007, 2008, 2011, 2015).

Opportunity for control is another category connected with the task conditions factors in TCF. As *opportunity for control* concerns how much influence task-takers can exert on task and the way it is implemented through such opportunities as extra time to complete the task, an easy task content (Skehan, 1998), asking clarification questions and negotiating meaning and the task itself (Skehan, 1996), it may directly relate to *planning time*, *few contributions needed* and *negotiation not needed* dimensions in TCF as these dimensions concern the whether or not time is given before and/or during task implementation (*planning time*), one or more individual contributions are required (*few contributions needed*), and no, little or extensive negotiation is required for contributing to the interaction during task implementation (*negotiation not needed*).

Clarity and sufficiency of information in *Task Difficulty Scheme* deals with the degree of directness with which task-related information is provided and the degree to which inference should be made. This category may relate to *task structure* in TCF since it examines if there is a clear structure regarding task steps.

Time pressure refers to the speed with which tasks are expected to be finished (Skehan, 1998, p. 142). It denotes the time limit within which a task is expected to be completed and may also involve the time given for both pre-task planning and online planning during task implementation. Therefore, time pressure may correspond to

planning time in TFC by Robinson, which concerns whether learners are provided with planning time or not (Robinson, 2007, 2008, Robinson and Gilabert, 2007).

Considering the scopes of each framework and how they conceptualize the factors that alter task demands, the framework by Robinson seems to provide a relatively more detailed and precise account of task complexity factors (Awwad, 2017). Since Robinson maintains that *task conditions* and *task difficulty* cannot be measured during the task design stage, only the task dimensions concerning cognitive/conceptual demands and performative/procedural demands can be considered to examine task complexity in textbooks. For this reason, a study can primarily draw on TCF by Robinson and examine the complexity of tasks with regard to two groups of task complexity variations: *resource-directing* and *resource-dispersing* dimensions. In addition, Skehan's conceptualization of *cognitive familiarity*, would be helpful for the conceptualization of *prior knowledge* in such a query. Two subdimensions of *cognitive familiarity*, *familiarity of topic* and *familiarity of discourse genre*, can be used to explain the boundaries of PK dimension.

The other dimensions in the Skehan's framework that have to do with task input or task conditions would not be considered when examining task complexity for two reasons. First, Robinson (2008) argues that task conditions mainly concern the situational setting and the conditions in which tasks are implemented rather than the intrinsic features of tasks. Second, Robinson finds sequencing decisions based on linguistic content of texts problematic (2001b). For him, the textual input to tasks determined based on some factors, including learners' language proficiency, concerns task content rather than task sequencing that needs to be manipulated based on the increase in the cognitive complexity dimensions only.

Empirical Studies on Tasks

Previous research on tasks in ELT literature generally centred around particular issues such as the conceptualization of tasks by language teachers, researchers and ELT coursebook designers; the distribution of activities (tasks and non-tasks) and task types in the coursebooks; as well as the task demands resulting from the task complexity dimensions.

As task is a controversial concept that has been defined differently in ELT, one of the focus points of previous research is how it is operationalized. Since the task definition by Ellis is somewhat ambiguous, some studies were devoted to further explaining the term (Erlam, 2016; Jaruteerapan, 2020; Lambert, 2018), which yielded contradictory accounts of it, thus adding to the confusion. Moreover, previous studies demonstrated how coursebook designers (Barrot, 2019) and language teachers differed in their understanding of the concept of task or the way they design tasks (Dao & Newton, 2021; East, 2018; Erlam, 2016; Jaruteerapan, 2020; Oliver & Bogachenko, 2018; Zhu, 2020). Previous research also revealed how and why language teachers transformed or replaced the tasks given in the textbooks (Nyungen et al., 2018). For instance, East investigated the design, implementation and evaluation of tasks by teachers and found out how “task-in-theory” could be somewhat different from “task-in-action” (2018, p. 41).

One instance is the study by Erlam (2016), which investigated how teachers conceptualized and designed tasks following the criteria proposed by Ellis (2003). The teachers received training mainly on TBLT for one year within the scope of a professional development program and designed tasks to be evaluated against the criteria in question. The analysis of these tasks revealed that more than half of the participating teachers still had difficulty understanding some aspects of tasks and

incorporating them into their task design. Similar findings were found in the study by Dao and Newton (2021) regarding the insufficient knowledge of language teachers of what the concept of task denotes and their tendency to reduce or replace communicative and task-based features of tasks while teaching.

Another contribution to TBLT was made by Jaruteerapan (2020), who modified the framework Erlam (2016) proposed to explain the task criteria by Ellis (2009). Arguing that the wording of the criteria by Ellis (2009) may have misled the teacher participants in Erlam's (2016) study to assume that tasks involve productive rather than receptive language, Jaruteerapan (2020) modified Erlam's framework to highlight the possibility of communication's being receptive or productive. Against this background, Jaruteerapan (2020) conducted an exploratory, qualitative case study to investigate EFL student teachers' experiences in TBLT training, in which they practised TBLT components through activities such as planning a task-based lesson for five weeks. The study showed that although most participants developed an understanding of TBLT principles and a positive attitude towards TBLT, they had difficulty grasping task features and constructing tasks. It was found that the participants particularly failed to understand and apply the *gap* (criterion 2) and *outcome* (criterion 4) principles.

Similar results regarding the misunderstanding of the task criteria were reached in the study by Peng and Pyper (2021). The study focused on the process of instructors' attempts at using tasks in a Chinese university context. It was discovered that several of the activities practised by the instructors and claimed to qualify for tasks were, in fact, task-like activities as they often failed to introduce an information gap to be filled with the task performance (criterion 2) and a non-linguistic task outcome to be accomplished as a result of this performance (criterion 4).

Another group of studies was devoted to investigating tasks in the ELT/EFL textbooks. The first issue to be discussed in this section is the frequency and percentage of tasks and non-tasks (i.e., exercises) in the textbooks. In the field of TBLT, it is maintained that workplans that target the use of explicit knowledge should involve learners in text-creation activities (i.e., tasks) rather than task-manipulation activities (i.e., exercises) for the development of implicit knowledge (Ellis, 2016). Nevertheless, ELT textbooks generally do not separate the use of explicit knowledge from text-manipulation activities, namely exercises (Ellis, 2016). It was discovered that the textbooks designed for older students predominantly consist of workplans focusing on text-manipulation rather than text-creation workplans (Ellis, 2016; Terekhova, 2017). They fall short in presenting sufficient input necessary for “the gradual and dynamic acquisition of implicit knowledge” and opportunities for social interaction to foster implicit knowledge and incidental learning (Ellis, 2016, p. 217). It was interpreted as the maintenance of the traditional exercise-based activities in the textbooks (Barrot, 2019; Ellis, 2016; Terekhova, 2017).

The place and role of tasks in ELT syllabus is another commonly investigated issue. Previous research revealed a limited number of task-based language textbooks (Ellis, 2018). Although some coursebooks were advertised as *task-based*, the textbooks examined in several studies were detected to be *task-supported* at the most (Ellis, 2018; Newton & Bui, 2017; Samuda & Bygate, 2008). Similarly, it is argued that most English textbooks adopt a task-supported approach to language teaching (Révész, n.d.).

Günay (2007) explored the tasks in an English textbook designed for 4th graders in state schools in Turkey. Within the scope of the study, a task criteria

checklist was developed, drawing on the task definitions by Carless (2003), Cameron (2001) and Brewster et al. (1992).). The criteria determined were (1) “Does the activity have a *clearly set input* that will be a starting point towards achieving the stated goal?”, (2) “Does it have an *outcome* that can be evaluated in terms of language production? OR Does it require the *use of language in the process of production*?”, (3) “Does it involve the learner in completing the activity for meaningful purpose *different from the displaying of the pre learnt language items*?”, and (4) “Does it involve the learner in a process of using *cognitive skills and psycho-motor skills at an optimum level of difficulty*?” (Günay, 2007, p. 103).

The study revealed that there was not a task-based progression throughout the textbook in question. The so-called tasks in the coursebook generally failed to satisfy the task characteristics commonly referred to in the TBLT literature. Although the coursebook had sections devoted explicitly to tasks, only a minority of these activities could meet the predetermined criteria to qualify as a task, thus indicating that how a task is perceived as a notion and presented in a textbook can vary. It is important to note that Günay (2007) comments that although most of the activities in the coursebook were categorized as non-task based on the criteria utilized, some of the non-tasks carry certain characteristics of tasks. The “insufficiency of structured input in their representation” was one of the reasons why Günay (2007) was not able to code some activities as tasks (p. 117).

Utilizing Günay’s (2007) checklist generated drawing on the task definitions by Carless (2003), Cameron (2001) and Brewster et al. (1992), Baklacı (2018) analysed the writing activities in the course materials used to teach Turkish as L2 in A1 and A2 levels. The analysis revealed that in both coursebooks most of the activities fail to qualify for task, mainly due to the fact that they fail to satisfy the

criterion 3, namely having a meaningful purpose “different from the displaying of the pre learnt language items” (Günay, 2007, p. 103).

Butler et al. (2018) investigated the activities presented in the English coursebooks for grades 3-6 in South Korea and China. The activities in the textbooks were grouped into three categories: “tasks”, “input-based tasks” and “non-tasks” (p. 286). Tasks in this study denote the activities that meet task criteria by Ellis (2003) and require language production. The analysis of the activities revealed that only a small portion of these activities would qualify as “task” or “input-based task”, whereas the majority of them were coded as “non-tasks” (Butler et al, 2018, p. 289).

A study by Barrot (2019) focused on task complexity in 30 ELT textbooks used in the Philippines, drawing on Ellis’s (2003) framework to distinguish tasks from exercises as the initial step of the study. The analysis of the activities revealed that the proportion of tasks in those coursebooks was relatively low vis-à-vis exercises.

Another study by Dao and Newton (2021) evaluated to what degree the activities in the coursebook, *New Cutting Edge* (Cunningham & Moor, 2005), were task-like and communicative based on the frameworks by Ellis (2008) and Littlewood (2004), respectively. Even if a considerable number of activities in the coursebook were found to qualify as task or task-like, these activities were discovered to be of low communicative value.

Terekhova (2017) conducted a study to compare the methodology of teaching English in Russia and Spain by analysing the activities in one textbook widely used in each country that targets the same proficiency level. The study evaluated the activities in the coursebooks (*Living English* and *Enjoy English*) against the suggestions of the Common European Framework of Reference (CEFR) for the

methods to be adopted for teaching English as a foreign language. Specifically, it aimed to arrive at a theoretical framework for the concept of task, reveal the common structures and activities in the coursebooks, categorise the activities in the first units of each coursebook and compare the quantity of the tasks in those coursebooks. The notion of task was conceptualized drawing on the definitions by Ellis, Richards, Willis, Nunan and British Council. Then, the activities in those coursebooks were categorised as exercise, task-like activity and task. It was discovered that most of the activities were designed as exercises focused on form. Interestingly, no tasks were found in one of the coursebooks. The other book included tasks, but their number was lower than task-like activities and exercises.

Another interesting finding of the study was that some activities in the coursebooks would easily qualify as task with a few modifications. For instance, it was found that the majority of the activities did not necessitate interaction. Although they satisfied other criteria to meet the requirements of taskness, these tasks were coded “task-like activity”, as they lack meaningful communication between students. This finding is worthy of consideration as it implies that coursebook designers and language teachers can easily transform task-like activities into tasks if taskness criteria are well explained.

The study by Rodríguez Sánchez (2018) was an attempt to offer a proposal for the task-based redesign of a synthetic structural coursebook to transform it into a task-supported one. The paper started by introducing the main components of TBLT, the approaches to teaching grammar and the organization of the language syllabus. Specifically, the notion of task, the three phases of task and task implementation process were defined. Then, grammar teaching approaches and the issue of focus-on-form were analysed. In what follows, approaches to syllabus and task design were

examined. Finally, a task-based redesign of the first unit of a coursebook, *Superminds*, was proposed to comply with the task criteria by Ellis and the TBLT model. The analysis of the textbook in the light of the relevant literature in the field of TBLT and the redesign process showed that it is feasible to integrate a TBLT approach into a language material originally designed with a structural Presentation-Practice-Production (PPP) approach through such strategies as shifting focus from form to meaning and adding a clear outcome.

Another issue commonly investigated is the task types. It has been observed that the majority of the studies in the field of TBLT have focused on *output-based* tasks (Ellis, 2017) and, specifically, oral production (Ellis, 2015, 2018; Sasayama, Malicka & Norris, 2015). So far, few studies have investigated *input-based* tasks (Erlam & Ellis, 2018; Shintani, 2012, 2014, 2016; Takimoto, 2009), and it has been observed that the ones investigating this type mainly focused on the activities that required learners to listen to someone and perform an action (Ellis, 2015).

The types of tasks analysed by researchers for task-based curriculum show differences. For instance, Willis (1996) and Long (1985, 2014) focused on *real-world* tasks, whereas Skehan (1998) investigated only *pedagogic production* tasks. On the other hand, Ellis (2003) studied both *real-world* and *pedagogic* tasks for TBLT.

As for the frequency and percentage of task types in the ELT textbooks, they were examined based on different task typologies. Some studies utilized Nunan's framework for coursebook evaluation (Alemi et al., 2013; Alemi & Hesami, 2014; Huda et al., 2020). One study that explored task types in an EFL textbook used in the Turkish context (Günay, 2007) by utilizing the framework by Willis (1996). Another

study (Papi, 2015) evaluated task types in an ELT coursebook based on the task type checklist developed by Finch (1999).

One common finding of the previous studies on task types in the ELT coursebooks is that there were discrepancies within and across the ELT coursebooks in terms of the frequency and distribution of task types as particular task types were given special importance (Alemi et al., 2013; Alemi & Hesami, 2014; Günay, 2007; Huda et al., 2020; Yoon, 2017). Such an unsystematic distribution of task types in the coursebooks would affect L2 development.

Günay (2007) aimed to analyse tasks in terms of their constitutive features, type variation, skill distribution and sequencing with specific respect to task demands. Task types were identified based on the framework by Willis (1996). Finally, a multi-factorial analysis of tasks was conducted to examine them in terms of “language, cognitive involvement and interactional demands” (Günay, 2007, p. 169). The study revealed that the task types and language skills focused were unevenly distributed, failing to support balanced language development. There was an emphasis on monologic tasks. Moreover, the analysis of the task sequence in terms of task demands demonstrated that linguistic complexity was gradually increased throughout the coursebooks whereas cognitive involvement and interaction demands were not observed to show a steady increase as units progressed.

Surveying the literature, the researcher observed that a limited number of studies investigated task types in the ELT textbooks based on the typology proposed by Ellis (2003, 2011). The studies mainly centred around dual categories by Ellis (2003, 2011) such as *focused* versus *unfocused* (Ahour & Shemshadsara, 2015; Alavinia et al., 2018), *input-based* versus *output-based* (Alavinia et al., 2018; Hamavandy & Golshan, 2015), *open* versus *closed* (Lee, 2020; Zohrabi &

Hassanpour, 2020), *one-way* versus *two-way* (Shadabi et al., 2017), and *real-world* versus *pedagogic* (Rahimpour & Kiani, 2015) to compare each type in terms of task-takers' motivation, performances (e.g., CAF) and language acquisition (e.g., grammar, vocabulary).

Nitta and Gardner (2005) focused exclusively on form-focused tasks, namely “grammar consciousness-raising tasks, interpretation tasks, focused communication tasks, grammar exercises, and grammar practice activities”, and analysed nine intermediate-level textbooks commonly used in the U.K. The first three types were conceptualised using the framework by Ellis (2003). Consciousness-raising tasks were found to be the most frequently presented tasks in those textbooks, followed by interpretation and focused communication tasks. The relative unpopularity of focused communication tasks was attributed to the difficulty of incorporating tasks that necessitate “presentation through production” into the traditional model that follows a different order (Nitta & Gardner, 2005, p. 10).

The final issue to be discussed in this section is the results of the previous research on task demands. So far, a body of research has been devoted to investigating how task characteristics influence L2 acquisition processes and how task demands lead to variation in L2 spoken and written production during task implementation. The studies show variety in terms of the factors they focused on, which have the potential to change the task performance. These factors include task type (Bistline-Bonilla, 2020; Cho, 2011; Gilabert, 2007; Vo, 2018), type of modes (Moradi & Farvardin, 2019; Smith, 2001), modality (Baralt, 2010), genre (Yoon, 2017), task sequencing (Hardy & Moore, 2004; Robinson, 2001b) and proficiency level (Cho, 2011; Riccardi, 2014; Yoon, 2017).

In parallel with the trend in the field of language teaching towards case studies to focus more on the situated nature of teaching and learning contexts (Motteram & Thomas, 2010), task complexity studies have also begun to employ micro perspectives (Révész, 2011). Although previous research primarily studied the task performance of language learners by examining their L2 production with CAF measures through experimental or semi-experimental research designs, studies in the last two decades – although their numbers are limited – have also focused on the task implementation process to examine the effect of task complexity, specifically on the interaction and negotiation (Cho, 2011; Gilabert et al., 2009; Hardy & Moore, 2004; Moradi & Farvardin, 2019; Vo, 2018) of learners. These studies drew on the claim of CH that increasing task complexity facilitates more opportunities for negotiation and interaction (Robinson, 2003, 2015) and, therefore, language learning.

Some attempts have been made to utilize previous frameworks when making judgments about task demands. For instance, Luo (2007) developed *Input-Processing-Output (IPO) Task Difficulty Matrix* to predict task difficulty at global and analytic levels based on the theoretical frameworks and empirical studies by Skehan (1996), Norris et al. (1998) and Brown et al. (2002). A greater number of studies, however, examined task complexity by drawing on the predictions of CH by Robinson (2001a, 2001b, 2003, 2005, 2007) and focused on manipulating the variables given in the TCF (Robinson, 2007).

Among the three factors defined in the model, task complexity (cognitive factors) has drawn more attention, while task conditions (interactive factors) and task difficulty (learner factors) were studied less. It can be because Robinson argues that neither task difficulty nor task conditions provide a solid base for sequencing pedagogical tasks as it is not always possible to have the knowledge of learner-

related factors or interactive factors to be involved during task implementation (2001a).

The first category of the TCF, *task complexity*, has two dimensions: *resource directing* and *resource dispersing* (Robinson, 2001a, 2001b). Among the task complexity features, a considerable focus was directed to the *resource-directing* variables. The resource-directing variables studied include *here and now* (Gilabert, 2007; Gilabert et al., 2009), *few elements* (Donate, 2018; Gilabert, 2007; Gilabert et al., 2009; Gökgöz, 2009; Lee, 2018; Révész, 2011; Riccardi, 2014; Sasayama, 2015) as well as *reasoning* (Awwad, 2017; Awwad et al., 2017; Choong, 2014; Gilabert, 2007; Gilabert et al., 2009; Révész, 2011; Riccardi, 2014).

The second group includes resource-dispersing (resource-depleting) variables that are associated with performative/procedural demands of tasks. The studies which explored these variables focused on *planning* (Hsu, 2012; Larson, 2020; Tabari, 2018; Yıldız, 2017; Yuan, 2001), *task structure* (Alwi, 2010; Alwi et al., 2012), *prior knowledge* (Adams & Alwi, 2014; Arslanyılmaz & Pedersen, 2010; Hardy & Moore, 2004; Larson, 2020; Robinson, 2001b) and *task support* (Alwi et al., 2012; Hardy & Moore, 2004).

The second category in the TCF, “task condition” dimensions, is related to “the interactive demands of pedagogic tasks” and depends on the nature of participation and participant variables (Robinson, 2001a, pp. 292-294). In previous studies, task condition factors were examined in terms of *interlocutor type* (Bistline-Bonilla, 2020) and *open solution variable* (Lee, 2018).

The existing research shows that cognitive task complexity was studied in qualitative and quantitative research designs and measured through various interaction measures for triangulation purposes such as L1 (first language) use,

repairs, recasts, self-corrections, question use, time estimations, interaction patterns, CAF, conversational negotiations, self-ratings of task difficulty, task-related episodes (TREs) and language-related episodes (LREs).

As for the task design based on the framework by Robinson, three studies are worthy of consideration. Within the scope of the study by Minas (2017), communicative and grammar-based tasks were designed and implemented drawing on the CH (Robinson, 2001b, 2005, 2010, 2011). The study (Minas, 2017, p. iii) aimed at investigating “core complexity features in task design”, “task-based interaction”, “task-based assessment” and “task sequencing” to provide solid theoretical evidence for task-based teaching of young learners of isiXhosa, taught as L2 in the Eastern Cape in South Africa, for young beginner level learners. The study examined the cognitive complexity of the tasks after manipulating the task design features based on the SSARC model with the aim of sequencing them to affect L2 (second language / foreign language) development (2017).

The other study by Un-udom and Patanasorn (2020) aimed to provide empirical evidence for TCF by investigating task complexity dimensions. With three tasks differentiated in their levels of complexity based on SSARC principles, the study assessed task complexity via three instruments: “a task difficulty assessment questionnaire, a report on time needed to complete the tasks, and a stimulated recall session” (p. 1). The results of the study confirmed the assumptions of the SSARC model by two measurements out of three – the task difficulty questionnaire and stimulated recall – thus proving that sequencing task complexity dimensions as described in TCF and recommended in the SSARC model would be beneficial for sequencing tasks in a hierarchical order of cognitive complexity.

Only two studies so far have utilized the framework by Robinson to investigate task complexity in ELT coursebooks (Barrot, 2019; Butler et al., 2018). Barrot (2019) examined the complexity of tasks within and across different grade levels (1 to 10) in 30 ELT textbooks used in the Philippines. First, four criteria for taskness proposed by Ellis (2003) were used to distinguish tasks from non-tasks. Then, drawing on the theoretical frameworks by Skehan (1998, 2003) and Robinson (2001b, 2003, 2005, 2007), Barrot (2019, pp. 471-472) created a task complexity matrix to judge task complexity levels and task sequencing, which consisted of “contextual support”, “task type”, “modality”, “task duality”, “complexity of steps”, “planning time”, “prior knowledge”, “control”, and “time pressure” as task complexity dimensions. The study revealed that less than one-fourth of the activities in these textbooks qualify for a task, and the ones that qualify for a task were not found to follow a linear progression from simple tasks to complex ones within and across grade levels.

The other study by Butler et al. (2018) investigated the tasks in the EFL textbooks designed for primary school context (grades 3-6) in China and Korea regarding the cognitive demands they made on learners and compared the differences across grade levels. For the cognitive dimensions to be examined, “here-and-now” (coded as “there-and-then”), “reasoning demand”, and “perspective-taking” were selected (p. 288). As for the *there-and-then* dimension (-HN), mixed results were achieved as four of the textbooks were found to consist of smaller proportions of these tasks, whereas it was not the case for two other textbooks examined. Regarding *reasoning dimensions* (causal and intentional) and *perspective-taking*, they were all found to be rather low in frequency and proportion. Moreover, their frequency and proportion were not found to increase in parallel with grade levels because the tasks

were manipulated based on linguistic elements rather than cognitive levels across grade levels.

All in all, this chapter presented a summary of the theoretical underpinnings of the present study. Starting with the concept of task and the introduction of TBLT in the first section, the chapter proceeded to the place and role of tasks in the textbooks designed in the field of ELT. The following sections focused on task characteristics and task types. Presenting the proposals for sequencing tasks, the fifth section examined the issues of task difficulty and task complexity. Finally, the last section of the chapter reviewed some of the relevant studies previously conducted in the field of ELT. The next chapter of the thesis will be devoted to the methodology of the study.

CHAPTER 3: METHOD

Introduction

This chapter explains the methodology that was followed in the present thesis. With this aim, the first section of the methodology chapter begins with a discussion of the research design with specific regard to the research problem, research questions, the research design, and data collection and analysis processes. The second section describes the context of the study, focusing on English language education policy, the secondary education curriculum, and the EFL textbooks used in Turkey. Then, it proceeds to the method followed for sampling. In what follows, the fifth section explains what instruments and data collection tools were utilised, explicating how their use fulfilled the purposes of the study. The next two sections briefly explain how the research data was gathered and analysed. Finally, the last two sections of the methodology chapter are devoted to the intercoder analysis and ethical considerations.

Research Design

This study was conducted as a response to the increasing focus on task-based approaches in the field of ELT (Motteram & Thomas, 2010) and the demand for research on the distinct characteristics of different tasks (Samuda & Bygate, 2008), task types as well as task complexity, as a critical consideration for syllabus design (Gilabert 2007; Robinson, 2007). Therefore, the present study seeks to examine the activities in the EFL textbooks used in the Anatolian High school context to reveal the task characteristics, task types and task complexity dimensions in these textbooks. The study also aims to demonstrate how the grade levels compare and

contrast with one another regarding these issues. In line with the research purposes, the following research questions were addressed in the present study:

1. To what extent are the activities in the English textbooks used in the Anatolian High school context possess task characteristics?
2. How are task types sequenced in the textbooks?
3. How are task complexity dimensions of output-based and integrative tasks sequenced in the textbooks?

The textbooks used in the present study were explored through content analysis, a method often used by qualitative researchers (Fraenkel & Wallen, 2009). Content analysis denotes “a family of research techniques for making systematic, credible, or valid and replicable inferences from texts and other forms of communication” (Drisko & Maschi, 2016, p. 7). This research method enables researchers to investigate the elements of almost all kinds of communication, including essays, articles and textbooks (Fraenkel & Wallen, 2009). Content analysis was determined to be employed in this study since it is accepted that “content analyses of course syllabi and textbooks are useful tools to document what is taught, and what is omitted, in specific educational programs” (Drisko & Maschi, 2016, p. 28).

Content analysis was preferred as the present study involves the quantification of qualitative data. Within the study, coding was performed both inductively and deductively. According to Fraenkel and Wallen (2009, p. 474), at the beginning of the content analysis, a researcher can decide the categories in line with the previous “knowledge”, “theory”, and “experience”. Likewise, a primary set of categories was deductively generated based on the theoretical framework of this study. However, the analysis of the contents of the textbook necessitated some new

categories and sub-categories. When data does not fit in the original, pre-determined categories or sub-categories, it is possible to generate new ones (Drisko & Maschi, 2016). Hence, inductively created categories or sub-categories were later added to the coding scheme during the analysis.

Context

English Language Education Policy

Turkey's current K12 education system was introduced with the curriculum reform of 2012, which extended compulsory education from eight to 12 years (2-8). K12 education is now divided into primary, secondary, and high school, each lasting four years.

With the new 4+4+4 tripartite system, English instruction started to be provided from grade 2 onwards rather than grade 4, as previously applied. The new system necessitated revising English language curriculum to address a wider age group with the addition of two new grade levels (2nd and 3rd). The new curricula for grade levels 2-8 (primary education) and 9-12 (secondary education) were designed to follow the principles and descriptors of the Common European Framework of Reference (CEFR) and to reflect the language proficiency levels (e.g., A1, B2) based on the level indicators of the framework (MoNE, 2018a).

The weekly course load for English shows variety in each education level. It increases as grade levels increase in primary and secondary education. While English course is provided for 2 hours in grades from 2 to 4, it is 3 hours in grades 5 and 6 and, finally, 4 hours in grade levels 7 and 8 (MoNE, 2018a).

As for secondary education, the number of weekly English lessons depends on the type of school. Students generally take English for General Purposes (EGP) for four hours in high school grades. However, they can also do additional English

courses offered in some types of high schools. For instance, the students in vocational and technical high schools can do compulsory and elective English courses in their field as well as the general English courses after choosing their departments in grade 10. Besides, there are high schools that offer preparatory classes, where students study English for 20 hours per week throughout the school year before starting grade 9.

In parallel with the increase in levels and grades throughout K12 education, language formality and complexity are increased. It is intended for students to start with A1 level (2nd-6th grades) and continue with level A2 (7th-8th grades). In the secondary level, 9th grade begins with A1/A2 level to revise and recycle the language content offered in the previous classes since students may start this grade with varying capacity and English proficiency levels as well as individual differences (MoNE, 2018b). It is also noted that the A1 level in this grade is relatively more complex in terms of language structures and vocabulary.

Since the English language curriculum for primary and secondary education addresses a broad age group (6-13), possessing different cognitive and social characteristics, the curricular model for primary education is divided into different stages. However, secondary education caters for a close range of learners, making a smooth transition from one grade to another possible with the instruction as a continuum without strict boundaries or stages.

Secondary Education Curriculum

According to the English Instructional Plan by the Ministry of National Education (MoNE), an “action-oriented approach” was adopted in the curriculum design process of the textbooks, and language was taken as a means for communication, not as an end (2018b, p. 4). *Action-oriented approach* was preferred

to enable students to act as a language user, practice language as a means of communication, and improve their communicative competence by emphasising functions and language skills in an integrated fashion.

The curriculum aims “to engage learners of English in stimulating, motivating, and enjoyable learning environments so that they (students) become effective, fluent, and self-directed users of English” (MoNE, 2018b, p. 4). Besides, collaborative activities are integrated into the curriculum to foster collaboration among students, which is likely to initiate interaction for negotiation of meaning and development of interpersonal skills. In addition, learner autonomy is intended to be supported through collaboration, interaction, communication, self-reflection, and decision-making activities. It is also aimed to enable students to get engaged in “task-based, collaborative, and project-based language activities” (p. 6). Finally, authentic assessment tools and techniques are included in the new curriculum to better reflect the language teaching and learning methods adopted. These techniques are also stated to reflect the real-life tasks that are likely to be encountered outside the school on daily basis.

The curriculum is claimed to be “multi-syllabus” as it shares components from different syllabus types, especially from “functional syllabus” and “skill-based syllabus” (MoNE, 2018b, p. 13). Thus, it is stated to adopt an eclectic approach.

The syllabus of the curriculum focuses on three areas of learning: *The Functions and the Useful Language, Language Skills and Learning Outcomes*, as well as *Suggested Materials and Tasks* for each unit (MoNE, 2018b). The first part, *Functions and Useful Language*, includes communicative functions and particular language units related to functions with their samples, which are built on one another and presented cyclically. This part provides specific structures and lexis for teachers

and material developers. Secondly, *Language Skills and Learning Outcomes* specifies four basic language skills and learning outcomes related to the functions and useful language targeting theme and task requirements. Besides, information regarding pronunciation is provided in this section. The final section, *Suggested Materials and Tasks*, displays suggested materials (e.g., recipes, blogs) and tasks (e.g., debates, role plays) to address the needs of the learner profiles at different levels of development.

It is stated that the materials and tasks provided in the curriculum are designed to have authenticity, cater for learners' real-life language needs and interests, appeal to multiple senses of learners via multimedia and instructional technology tools, enable the use and assessment of four language skills in an integrated way in line with the principles of the Common European Framework of Reference (CEFR) and present characters and places familiar to learners across different units and grades to ensure permanent learning.

The textbooks were prepared in line with the updates in the English curriculum for 2nd-8th grades in 2013 and are considered as a continuum of the books used for those grades (MoNE, 2018b). They aim to develop four language skills and foster communicative competence, collaboration, learner autonomy and reflective thinking. With this aim in mind, the activities in the books were designed to be “task-based, collaborative, and project based” (MoNE, 2018b, p. 6). The Instructional Model by MoNE (2018b, p. 7) for each textbook is presented in Table 6.

Table 6*Model English Language (9th – 12th Grades) Curriculum*

Grades (CEFR*) Hours per Week	Learner Age	Skill / Grammar / Vocabulary / Pronunciation / Focus	Main Activities (Can be used in all grades)
9 (A1/A2) 4	14-14.5	All four skills integrated with an emphasis on Listening and Speaking Limited focus on language structures Maximum seven new Vocabulary items per lesson Limited Pronunciation practice	Roleplays/Simulations Graphics/Charts Paragraph Reading and Writing
10 (A2+/B1) 4	15-15.5	All four skills integrated with an emphasis on Listening and Speaking Maximum seven new Vocabulary items per lesson Limited Pronunciation practice	Information Gap Activities/Roleplays Paragraph Reading and Writing Skimming and Scanning
11 (B1+/B2) 4	16-16.5	All four skills integrated with an emphasis on Listening and Speaking Maximum seven new Vocabulary items per lesson Limited Pronunciation practice	Surveys/Short Oral Presentations/Drama Short Reading Texts and Paragraph Writing
12 (B2+) 4	17-17.5	All four skills integrated with an emphasis on Listening and Speaking Synthesis of Language Structures Maximum seven new Vocabulary items per lesson Limited Pronunciation practice	Conversations/Oral Presentations/Projects/ Task-based Activities Argumentative/Descriptive Text Writing

Note. From *Ortaöğretim İngilizce Dersi (9., 10., 11. ve 12. Sınıflar) İngilizce Öğretim Planı* (p. 7), by Ministry of National Education (MoNE), 2018, Milli Eğitim Bakanlığı. Copyright 2018 by Milli Eğitim Bakanlığı.

It is seen in Table 6 that the objectives and outcomes of the textbooks were determined based on descriptive and pedagogical foundations of CEFR for learner profiles. Each grade level starts from the exit level of its preceding level. The CEFR levels of the textbooks in the following grades incrementally increase, beginning with the A1/2 (Basic Users) in the first textbook prepared for 9th graders, proceeding to A2+/B1, B1+/B2 and finally targeting the B2+ level (Independent Users) in the fourth book prepared for 12th graders. With this aim in mind, each textbook was

designed to consist of activities geared toward improving four language skills, grammar, vocabulary and pronunciation.

EFL Textbooks Used in Secondary Education

In Turkey, secondary education lasts four years and serves the learner groups roughly between the ages of 14-17. The education services and learning materials, including textbooks, are financed by the state in public schools. The textbooks within the present study were designed by a local community in the Ministry of National Education (MoNE), targeting the education system in the high schools in Turkey. Therefore, they are geared towards the ages, language levels and language needs of the target learner groups.

Each textbook comprises ten themes to be completed within a weekly 4-hour course load during two semesters in one education year (MoNE, 2018b). The organisation of the units in the textbooks examined in the current study is presented in Table 7 (Akdağ et al., 2019, pp. 8-9; Bulut et al., 2019, pp. 8-9; Çimen et al., 2019a, pp. 8-9; Çimen et al., 2019b, pp. 8-9).

Table 7

The Organisation within Units in the High School EFL Textbooks

Order within Units	Grade Level / Textbook			
	9 Teenwise	10 Count me in	11 Silver Lining	12 Count me in
1	Theme	Theme	Theme	Theme
2	Function	Reading	Function	Reading
3	Grammar	Listening	Grammar	Listening
4	Vocabulary	Speaking	Vocabulary	Speaking
5	Listening and Reading	Writing	Listening and Reading	Writing
6	Speaking and Writing	Punctuation	Speaking and Writing	Punctuation
7	Punctuation	Sample Usage	Punctuation	Sample Usage

In Table 7, it is demonstrated that four language skills and pronunciation are emphasised throughout the curriculum. Each unit in all the textbooks starts with the introduction of the theme and is followed by language skills, pronunciation as well as grammar (in grades 9 and 11) or sample usage (in grades 10 and 12). It is seen that the organising principles of the textbooks for grades 9 and 11, and grades 10 and 12 are the same. It can be due to the fact that the textbooks for grades 9 and 11, and grades 10 and 12 were written by almost the same groups of writers.

Sampling

Four EFL textbooks used in the Anatolian high schools in Turkey during the 2020-2021 education year were chosen to explore task characteristics, task types and task complexity. The selection of the textbooks for analysis within the present study was purposeful. As secondary education lasts four years in Turkey, it was aimed to select one textbook for each grade. Only the textbooks used in the 2020-2021 education year were examined to gain insights into the current condition of task design in EFL textbooks.

As the study aims to reach more representative data on the design and implementation conditions of pedagogical language tasks in the high school context in Turkey, it necessitated determining which textbooks are used by a higher number of high school students. Since Anatolian high schools outnumber other school types at the secondary level in Turkey (MoNE, 2021), the textbooks used in this type of school were considered. These schools have two options for the textbooks to use for English courses, one offered by commercial publishing houses and the other by MoNE. Among them, only the textbooks designed by MoNE designed in line with the English language education policy and secondary education curriculum were preferred. It was considered that the analysis of the English textbooks designed by

MoNE to be used in the Anatolian high school context would yield more representative results regarding the current condition of language tasks in ELT textbooks in Turkey.

As a result, four textbooks, *Teenwise* (A1/A2), *Count me in* (A2+/B1), *Silver Lining* (B1+/B2) and *Count me in* (B2+...), were chosen to analyse tasks in the EFL textbooks for the 9th, 10th, 11th and 12th grades, respectively.

Instrumentation

All the coursebooks can be accessed from the website of the Educational Informatics Network (EBA) provided by the MoNE. Microsoft Excel spreadsheet software was used to code and categorise and sift and sort the data collected from the textbooks.

Checklists

To address the research questions, three checklists were developed by the researcher based on the theoretical frameworks used in this study (Table 8).

Table 8

Frameworks Used for the Development of the Checklists in the Analysis Stages

Checklist	Focus Points	Frameworks
1	Task characteristics	Ellis (2009), Ellis and Shintani (2014) and Jaruteerapan (2020)
2	Task types	Ellis (2003, 2012, 2020)
3	Task complexity	Robinson (2001a, 2001b, 2003, 2007, 2008, 2011, 2015) and Skehan and Foster (2001)

Based on each framework given in Table 8, a set of categories and codes was generated deductively. When the data in the textbooks necessitated the generation of new categories or sub-categories, they were added to the checklists after the researcher and the other coders agreed. These categories were marked with the

asterisk symbol (*) in the instruments. Colouring was also preferred in the checklists to easily distinguish the categories used to work on language activities and tasks.

Checklist 1: Task Characteristics

The first checklist (Figure 3) was generated to be able to answer the first research question of the present study, which is “To what extent are the activities in the English textbooks used in the Anatolian High school context possess the task characteristics?”.

Figure 3

Checklist 1 Used for Identifying Task Characteristics

No	G	U	AN	A	Q1.1	Q1.2	C1	Q2.1	Q2.2	C2	Q3.1	Q3.2	C3	Q4.1	Q4.2	C4	D
1																	
2																	
3																	
4																	
5																	

Note. G = grade, U = unit, AN = activity number, A = activity, Q = question, C = criterion, D = decision on taskness.

The checklist above (Figure 3) was drawn on the task characteristics discussed by Ellis (2009), Ellis and Shintani (2014) and Jaruteerapan (2020) to distinguish tasks from non-tasks. The checklist consists of columns for the number of the activities (as assigned by the researcher), grade, unit, activity number (as specified in the textbooks), activity, task criteria (C1-C4), as well as the decision on taskness (i.e., task vs non-task).

The four main items (C1, C2, C3, C4) in the checklist refer to the four task characteristics discussed by Ellis (2009, p. 223) and Ellis and Shintani (2014, pp. 135-136), which are (1) “The primary focus should be on ‘meaning’”, (2) “There should be some kind of ‘gap’”, (3) “Learners should largely rely on their own

resources” and (4) “There is a clearly defined outcome other than the use of language”.

As it was observed in the previous studies that the four criteria (C1-C4) in question can be ambiguous for language professionals (Ellis, 2018; Erlam, 2016; Jaruteerapan, 2021), the coding framework by Jaruteerapan (2020) was utilised to explicate each criterion so that more accurate decisions on the taskness of the activities can be made (Table 9). The sub-categories for each criterion (e.g., Q1.1, Q1.2 etc.) in Checklist 1, denote the questions created by Jaruteerapan (2020).

Table 9

The Framework for Coding Activities Based on the Study by Jaruteerapan (2020)

Criteria	Questions	Required answer
1. A primary focus is on meaning. (C1)	Q1.1 Do the learners mainly use the language for a communicative purpose, not just learn the language (e.g., forms and functions) intentionally?	Yes
	Q1.2 Is the learner primarily concerned with producing and/or comprehending the messages (semantic and pragmatic meaning), not with focusing on linguistic form?	Yes
2. There should be some kind of gap. (C2)	Q2.1 Does the activity have a gap that requires a transfer of given message from one person to another, or one form to another (written to oral) or one place to another (texts to table) in order to close it?	Yes
	Q2.2 Is the gap closed as a result of learners’ participation in communication (i.e., as either as a speaker/writer or listener/reader)?	Yes
3. Learners rely on their own resources. (C3)	Q3.1 Are the learners not “taught” * the language they will need to perform the activity? * <i>Being taught the language does not include the teachers providing some linguistic starting point such as key vocabulary, an input for production tasks (priming)</i>	Yes
	Q3.2 Does the activity allow the learners to use the language they already knew to complete the activity?	Yes

Table 9 (cont'd)

The Framework for Coding Activities Based on the Study by Jaruteerapan (2020)

4. There is a clearly defined communicative outcome. (C4)	Q4.1 Are the learners primarily concerned with achieving the goal stipulated by the activity, rather than using language forms correctly?	Yes
	Q4.2 Is there an outcome that results from completing an activity that works towards a communicative goal, other than the display of linguistic knowledge?	Yes

Note. Adapted from *The Emerging Understandings and Practices of Task-Based Language Teaching (TBLT)* by Thai EFL Student Teachers (pp. 71-72), by P.

Jaruteerapan, 2020, [Unpublished doctoral dissertation]. Victoria University of Wellington. Copyright 2020 by Victoria University of Wellington.

It is seen in Table 9 that Jaruteerapan (2020) developed two conditions that should be satisfied for an activity to meet Ellis's (2009) each criterion of taskness.

For the first criterion (i.e., meaning), learners are expected to (1) utilise language mainly for a communicative purpose, not only to learn it and (2) focus mainly on the production or comprehension of messages, not on the linguistic form during task implementation.

For the second criterion (i.e., gap) to be satisfied, (1) a task should have a gap that necessitates "a transfer of given message from one person to another, or one form to another (written to oral) or one place to another (texts to table) in order to close it" and (2) the gap should be closed when learners participate in communication (Jaruteerapan, 2020, p. 71).

When it comes to the third criterion (i.e., resources), learners are (1) not taught the language required to accomplish the task and (2) allowed to use the language they already know (Jaruteerapan, 2020).

The final criterion (i.e., outcome) is also explicated with two conditions: (1) learners mainly focus on achieving the task goal instead of correct use of language forms and (2) the task outcome serves a communicative goal, not the display of linguistic knowledge (Jaruteerapan, 2020).

Notice that each question from 1.1. to 4.1 in Jaruteerapan's (2020) coding framework is present in the checklist used in the present study (Figure 3). For instance, 1.1 and 1.2 were named Q1.1 and Q1.2, respectively, to make it clear to which criterion they belong to (e.g., C1, C2, C3 or C4) and their order (1 or 2).

Checklist 2: Task Types

The second checklist (Figure 4) was created to find an answer to the second question of the present study, which is "How are the task types sequenced in the textbooks?". Hence, the checklist was geared toward determining the task types within the textbooks.

Figure 4

Checklist 2 Used for Determining Task Types

No	G	U	T	FO	UN	RW	PE	IB	OB	*INT	HN	TT	CL	OP

Note. G = grade, U = unit, T = task, FO = focused, UN = unfocused, RW = real-world, PE = pedagogic, IB = input-based, OB = output-based, *INT = integrative, HN = here-and-now, TT = there-and-then, CL = closed, OP = open.

Checklist 2 (Figure 4) was drawn from the classification made by Ellis (2003, 2012, 2020). For this checklist, items were prepared to classify the tasks as (1) *focused* (FO) versus *unfocused* (UN), (2) *real-world* (RW) versus *pedagogic* (PE), (3) *input-based* (IB), *output-based* (OB) versus *integrative* (INT), (4) *here-and-now*

(HN) versus *there-and-then* (TT), and (5) *open* (OP) versus *closed* (CL). To distinguish between each binary category easily, the columns for each dual category were coloured with a different colour.

It should be noted that although Ellis's (2003, 2012, 2020) task classification is based on dual categories, *integrative* (INT) tasks category was added as a third category to *input-based* and *output-based* tasks. Since this category was included in Checklist 2 during the task analysis process, it was marked with an asterisk sign (*) to make it clear that it was an addition.

It should be noted that the conceptualisation of INT tasks in the literature (Ellis, 2009, 2011, 2012) needs further explanation. It is unclear whether it denotes the use of more than one language skill (whether IB or OB) or both IB and OB tasks for implementing one task. For instance, a task may require two skills from the same task type, such as listening and reading (input-based) or skills from two distinct task types (IB vs. OB), such as reading and speaking. Since the relevant literature does not focus on the relevant difficulty levels of skills in the same category (e.g., listening versus reading) but those of the skills in different types (IB vs. OB), the researcher decided to operationalise INT tasks to refer only to the tasks that require the use of skills from both IB and OB tasks for the present study.

Checklist 3: Task Complexity

Checklist 3 (Figure 5) was designed to answer the third question of this study, which is "How are the task complexity dimensions of the output-based and integrative tasks sequenced in the textbooks?". The relevant checklist provided below serves to explore the complexity dimensions of the tasks in the textbooks.

Figure 5*Checklist 3 Used for Exploring Task Complexity*

No	G	U	T	Resource-directing Dimensions								Resource-dispersing Dimensions								SSARC Steps				
				(+) TT		(+) SR		CR	IR	PT	(-) ST		(-) TS		SS	(-) PK				Step 1 (-) *Rdir (-) *Rdis	Step 2 (+) *Rdis	Step 3 (+) *Rdir (+) *Rdis		
				(+) *SPA	(+) *TIM	(+) *STA	(+) *DYN				(-) *COS	(+) *MUL	(-) *CON	(-) *STE		(-) *TOP	(-) *DIG	(-) *CCD						

Note. G = grade, U = unit, T = task, TT = there-and-then, *SPA = space, *TIM = time, SR = spatial reasoning, *STA = static, *DYN = dynamic, CR = causal reasoning, IR = intentional reasoning, PR = perspective-taking, ST = single task, *COS = conceptual support, *MUL = multitasking, TS = task structure, *CON = context, *STE = steps, *SS = single step, PK = prior knowledge, *TOP = topic, *DIG = discourse genre, *CCD = cognitive/contextual demands, *RDir = resource-directing, *RDis = resource-dispersing.

Checklist 3 above (Figure 5) was generated drawing on the task complexity dimensions in the *Triadic Componential Framework* (TCF) by Robinson (2001a, 2001b, 2003, 2007, 2008, 2011, 2015) and the *Task Difficulty Scheme* by Skehan and Foster (2001). It consisted of two main categories of resource-directing dimensions and resource-dispersing dimensions. Resource-directing dimensions consist of *here-and-now* (HN), *spatial reasoning* (SR), *causal reasoning* (CR), *intentional reasoning* (IR) and *perspective-taking* (PR), whereas resource-dispersing categories comprise *single task* (ST), *task structure* (TS), *single step* (SS) and *prior knowledge* (PK).

Please notice that the resource-directing dimensions are accompanied by the (+) symbol, whereas the resource-dispersing dimensions by the (-) symbol to show increasing task demands. This distinction was made purposefully since while the presence of the resource-directing dimensions increases task complexity (in terms of

cognitive/conceptual demands), the absence of the resource-dispersing dimensions produces the same effect (in terms of performative/procedural demands). It should also be noted that the TT dimension (-HN) was used in the checklist for coding instead of HN so that all the columns for resource-dispersing dimensions in the checklist have the same symbol (+) and indicate the presence of a factor that increases the cognitive/conceptual demands tasks make.

Finally, to easily distinguish between resource-directing dimensions, resource-dispersing dimensions and task complexity steps based on SSARC from one another, the columns for categories were coloured differently.

The first category in the Checklist 3 is resource-directing dimensions. These dimensions refer to the factors that make “resource demands which can be met by using specific features of the language code” during task performance (Robinson, 2001b, p. 31). Five resource-directing factors examined in the checklist are *there-and-then* (+TT), *spatial reasoning* (+SR), *causal reasoning* (+CR), *intentional reasoning* (+IR) as well as *perspective-taking* (+PT). These dimensions were coloured in green to distinguish them from resource-dispersing dimensions, which pose different demands on learners when implementing a task.

The first sub-category of the *resource directing* dimensions was named TT (-HN) as it was aimed to represent each condition in the resource-directing dimensions in Checklist 3 with (+). Since the +HN condition stands for a simple task, it would not be appropriate to represent complex tasks in terms of cognitive/conceptual complexity. Therefore, its opposite, +TT condition, was preferred. TT condition is associated with specific “cognitive operations, conversational abilities, and linguistic resources” (Robinson, 1995, p. 102) needed for displaced reference to time and space and, therefore, directs attention to form-function mappings in the L2 during task

implementation (Robinson, 2014). Hence this condition was examined specially to find out if a task necessitates displaced references to (1) space and (2) time, increasing cognitive/conceptual complexity.

It should be noted that when +TT (-HN) condition was examined in the present study, it was considered independently from the *contextual support* (COS) condition, which is often considered within the scope of the HN condition in previous studies (Robinson, 1995, 2010, 2014). Nevertheless, considering the argument by Malicka (2014b) and Révész (2007, 2009) that HN should be regarded as a resource directing condition, whereas *contextual support* as a resource dispersing condition, only the HN condition promoting cognitive/conceptual complexity was considered independently from *contextual support*, posing procedural/performative demands on learners as resource-dispersing dimensions do.

The second resource-directing dimension, *spatial reasoning* (+SR), refers to the tasks that induce the employment of linguistic forms to describe motion events (Robinson, 2011). However, it was discovered that there are tasks in the textbooks that may require learners to deal with spatial relationships between objects, if not in motion events. Since previous studies have examples of tasks manipulated in terms of spatial reasoning, which deal with spatial reference to location (Levkina & Gilabert, 2014; Noevo, 2006; York, 2019), not motion events, the researcher decided to examine the tasks requiring reference to location as well. When coding the tasks, *static* (STA) and *dynamic* (DYN) categories were generated to differentiate between the tasks that involve location (STA) or motion (DYN).

The third resource-directing dimension, *causal reasoning* (+CR), relates to the tasks that necessitate reasoning about causal events and relationships between them (Robinson, 2007, 2008; Robinson & Gilabert, 2007).

The *intentional reasoning* (+IR) dimension refers to the tasks that require learners to reason about the motives, intentions, beliefs or desires of people as well as the relationships between these (Robinson, 2007, 2008, 2011).

The fifth and final resource-directing dimension, *perspective-taking* (+PT), deals with whether learners assume only one first-person perspective on an event (-PT) or several second and third-person perspectives to be able to accomplish a task (+PT) (Robinson, 2007, 2008; Robinson & Gilabert, 2007).

In addition to the *resource-directing* dimensions mentioned above, the second category in Checklist 3 is *resource-dispersing* dimensions, which make “demands on attention and working memory” during task implementation (Robinson, 2001b, p. 31). The resource-dispersing factors examined in the checklist are *single task* (-ST), *task structure* (-TS), *single step* (-SS) and *prior knowledge* (PK). They were coloured in yellow in the checklist to distinguish them from resource-directing dimensions.

The first resource-dispersing dimension in Checklist 3 is *single task* (-ST). Drawing from Robinson’s framework (2007, 2008), the tasks that necessitate only one thing to be completed are coded as +ST, whereas the ones that necessitate two or more things to be dealt with simultaneously are considered as -ST (*dual task*) within the present study.

It was argued in previous studies (Malicka, 2014b; Révész, 2007, 2009; Robinson, 2010, 2015) that the absence of contextual support may pose dual demands on task-takers as they need to both remember the details of an issue and try to accomplish task outcome. Therefore, one sub-category, *contextual support*,

(-COS) was added under the ST dimension and the tasks that lack contextual support were marked as -ST. For the other tasks that require learners to perform two actions concurrently for their implementation (e.g., listen to write, speak and draw),

another sub-category, *multitasking* (+MUL), was added to examine the ST dimension.

The second resource-dispersing dimension is *task structure* (-TS), which examines whether there is a clear structure provided so that learners can find out the contents of a task and distinguish the *task steps* (+TS) or not (-TS) (Robinson, 2007, 2008; Robinson & Gilabert, 2007). When a task was not presented with such *content* (-CON) as a model, template, sub-questions, or clear instructions for task steps (-STE) to be followed that would work towards guiding task-takers to implement the task to accomplish the outcomes, it was coded as lacking task structure (-TS).

When analysing *task structure* (-TS) of tasks, attention must be paid not to mistake the sub-category of *content* (-CON) for *contextual support* (-COS), which is a sub-category of the *single task* (-ST) dimension. While COS deals with whether task-takers access to the task content (e.g., visuals, task input) to be worked on during the task implementation, CON denotes the provision of a model, template and sub-questions that would help them to grasp the expectations as to the task outcomes.

Another issue to be careful about when interpreting task structure is *task steps* (STE). Some tasks consist of only one step and do not necessitate detailed instructions. In that case, only the CON sub-category was considered.

The third resource-dispersing dimension is the *single step* (-SS). Please note that this dimension is originally named *few steps* (FS) in Robinson's framework (2007, 2008; Robinson & Gilabert, 2007). Considering that no single operationalisation of FS was made to distinguish +FS condition from -FS (i.e., 'many steps') in the relevant previous literature (Robinson, 2007, 2008; Robinson & Gilabert, 2007), the researcher decided to reconceptualise this dimension as *single step* for the current study to refer to whether one step is involved in a task (+SS) or

more than one step (-SS) to accomplish the task outcomes. For the analysis of tasks with regard to the SS dimension, it is worthy of noting that when a task was marked as having more than one step (-SS), it must be interpreted that these steps are dependent on one another for the accomplishment of task goals.

The final resource-dispersing dimension examined was *prior knowledge* (-PK). Although Robinson's framework was taken as a base, in the current study *cognitive familiarity* dimension in the framework by Skehan and Foster (2001) was also utilised to operationalise the dimension.

The sub-categories of "cognitive familiarity" are "familiarity of topic", "familiarity of discourse genre" and "familiarity of task" (Skehan & Foster, 2001, p. 195). The first two sub-categories were integrated into the present study as familiarity of *topic* (TOP) and *discourse genre* (DIG). Instead of *familiarity of task*, however, the researcher added a new category that is concerned with *cognitive/conceptual demands* (CCD) as the third sub-category of PK dimension for two reasons. First, *familiarity of task* would overlap with the familiarity of *topic* and *discourse genre*, thus leading to "double-counting" of the overlap area, which is suggested to be avoided (Scriven, 2005, p. 3). Secondly, since the increased complexity in resource-dispersing dimensions is accepted to increase the cognitive/conceptual demands of tasks and learners' familiarity with these demands is predicted to affect performative/procedural demands of tasks (Robinson, 2001a, 2001b, 2003, 2007, 2008, 2011, 2015), the researcher decided to include the *cognitive/conceptual demands* as the third sub-category for PK dimension.

The last section of the Checklist 3 is devoted to the SSARC steps that demonstrate the complexity levels of tasks. SSARC model consists of three steps for task sequencing. Step 1 (-RDir -RDis) represents the complexity level in which both

resource-directing and resource-dispersing dimensions are simple. Step 2 (-RDir +RDis) shows that a task is made complex along resource-dispersing dimension(s) only. Step 3 (+RDir +RDis) indicates that a task is made complex along both dimensions. Please note that a new, intermediary step (+RDir -RDis) was incorporated into the checklist between the Step 2 and Step 3 to indicate the tasks that are made complex along the resource-directing dimension(s) only to represent the tasks with this complexity condition.

Method of Data Collection and Analysis

All the textbooks evaluated in the present study were publicly available on the website of the Educational Informatics Network (EBA) provided by the MoNE, where the soft copies of the textbooks were downloaded and saved for data analysis (2021, March 21).

This section explicates how data analysis was performed within the present study. Once the preliminary categories and sub-categories were determined, the focus areas of the study and the analysis stages were finalised (Table 10).

Table 10

The Overview of the Analysis Stages Based on the Research Questions

Research Question	Analysis Stages
1	Identifying task characteristics that the activities in the textbooks possess
2	Determining task types
3	Exploring task complexity dimensions of output-based and integrative tasks

As the first question of the current study (Q1) focuses on the extent to which the activities in the EFL textbooks used in the Anatolian High school context possess task characteristics, the first stage was to identify task characteristics the activities in the textbooks possess. Secondly, the types of tasks were determined since the second

question of the study (Q2) concerns how task types are sequenced in the textbooks. Finally, the task complexity dimensions of the output-based and integrative tasks were examined to answer the last question of the study (Q3), which deals with how task complexity dimensions of output-based and integrative tasks are sequenced in the textbooks. At each step of the analysis, grade levels were compared and contrasted with one another based on the issues analysed.

Analysis of Activities for Task Characteristics

The first question of the thesis is, “To what extent are the activities in the English textbooks used in the Anatolian High school context possess the task characteristics?”. It necessitated determining which activities could be considered a task or non-task (i.e., exercise). To that end, a checklist based on the task characteristics by Ellis (2009), Ellis and Shintani (2014) and the coding framework of Jaruteerapan (2020) was utilised.

The multifaceted nature of *task* defined by Ellis (2003) makes it difficult to operationalise the concept (Cobb, 2010). Previous studies (Butler et al., 2018; Erlam, 2016; Jaruteerapan, 2020) that drew on the different versions of the framework by Ellis (2003, 2009) regarding task criteria showed that these criteria still could be vague when it comes to making judgments on the *taskness* of the language activities and necessitate further conceptualisation. Therefore, the coding framework, which was designed by Jaruteerapan (2020) based on the task criteria by Ellis (2009) and the explanation of these criteria by Erlam (2016), was utilised to identify the task characteristics in the activities in the textbooks examined. In this way, Checklist 1 was developed to include eight questions for four criteria in total. A sample unit analysis for the identification of task characteristics is given below (Figure 6).

Figure 6

A Sample Unit Analysis for Task Characteristics

No	G	U	AN	A	Q1.1	Q1.2	C1	Q2.1	Q2.2	C2	Q3.1	Q3.2	C3	Q4.1	Q4.2	C4	D
320	11	10	a-b	a. Read the following statements. Guess and write true (T) or false (F). b. Read the text and check your answers.	1	1	1	1	1	1	1	1	1	1	0	0	0
321	11	10	1a	Look at the book covers. Which ones are familiar to you? Share and discuss what you know about the main characters.	1	1	1	1	1	1	1	1	1	1	0	0	0
322	11	10	1b	Read the poem and choose the best option to find the gist.	1	1	1	0	0	0	1	1	1	1	0	0	0
323	11	10	2	Read the sentences and choose the correct option for the words in bold.	0	0	0	0	0	0	0	0	0	0	0	0	0
324	11	10	3	Listen to Tanita and Kevin talking about the book The Little Prince and circle the correct option.	0	0	0	1	1	1	1	1	1	0	0	0	0
325	11	10	4	Listen to the dialogue again and write true (T) or false (F). Correct the false ones.	1	1	1	1	1	1	1	1	1	1	0	0	0
326	11	10	5a	Work in pairs. Look at some of the typical behaviours of a grown-up who feels young at heart. Brainstorm and add as many behaviours as you can to the list.	1	1	1	1	1	1	1	1	1	1	1	1	1
327	11	10	5b	Think of a grown-up who is young at heart from your family members / teachers / neighbours, etc. Then describe him / her.	1	1	1	0	0	0	1	1	1	1	0	0	0
328	11	10	1a-b	a. Read the quotes of The Little Prince and tick the ones you agree. b. Work in pairs. Compare your choices with your partner's and discuss.	1	1	1	1	1	1	1	1	1	1	1	1	1
329	11	10	2	Read the summary of The Little Prince focusing on the highlighted words. Use your dictionaries to choose their synonyms.	0	0	0	0	0	0	0	0	0	0	0	0	0
330	11	10	3	Read the summary again and answer the questions.	1	1	1	0	0	0	1	1	1	1	0	0	0
331	11	10	4	Read again to choose the main idea of the summary.	1	1	1	0	0	0	1	1	1	1	0	0	0
332	11	10	5	Work in groups. Write slogans on moral and social values like kindness, honesty, love, justice, etc. Create your own slogan on a T-shirt / mug / cushion / badge / backpack, etc.	1	1	1	1	1	1	1	1	1	1	1	1	1
333	11	10	a	There are two different pronunciations of either and neither. Listen and practise.	0	0	0	0	0	0	0	0	0	0	0	0	0
334	11	10	b	Listen and tick the sounds in either and neither words you hear.	0	0	0	1	1	1	1	1	1	0	0	0	0

Note. G = grade, U = unit, AN = activity number, A = activity, Q = question, C = criterion, D = decision on taskness.

As it can be seen in Figure 6 above, all the activities within four textbooks were entered in an Excel spreadsheet with relevant information as to the number of activities (No) as assigned by the researcher, grade, unit, activity number (AN), activity (A) as specified in the textbooks, task criteria (i.e., C1, C2, C3, C4), questions asked for each criterion (e.g., Q1.1, Q1.2) as well as the decision on taskness (D). If an activity was not assigned an activity number, then the initials of the activity heading were used in the activity number section (e.g., LE for *Let's Explore*).

It was necessary for an activity in the textbooks to meet four task criteria (with eight questions) to qualify as a task. Each criterion has two yes/no questions specifically developed to explain their nature. The number 1 was utilised to indicate that the answer to a question was *Yes* and 0 to indicate *No* to mark the checklist based on the task criteria with the questions given. Then, a decision was made on whether an activity can be considered a *task* or *non-task*. As for the decision on taskness, the column named *D* was utilised. On the condition that an activity was marked with 1 for each criterion (from C1 to C4) as well as the relevant questions (from Q1.1. to Q4.2) for the task criteria, this activity was considered as a task, and the section *D* was marked with 1 to indicate that it is a *task*. Otherwise, if an activity fails to satisfy any criteria, it was considered a *non-task* and marked with 0 under the same section.

As the study involves quantifying the activities in the textbooks, it necessitated deciding how to count these activities. It was discovered that there wasnot always a consistent format to assign numbers or letters to each activity in the textbooks to demonstrate that they denote an independent section or sections linked to each other. Therefore, the researcher manually counted each activity separately to

arrive at a more representative number. However, since some tasks require students to perform a few subsequent activities to be able to complete a task, and the number of steps within a task affects the task demands investigated in the present study, each task was counted only once within this study, no matter if it involves one step or more. The following sections explicate how coding was performed with regard to each criterion specifically.

Criterion 1

The first criterion requires that the primary focus is on *meaning* (Ellis, 2009; Ellis & Shintani, 2014). It requires that task takers mainly deal with the “semantic and pragmatic meaning” (Ellis, 2011, p. 212) and are engaged in encoding or decoding messages for communication (Ellis & Shintani, 2014). Two questions for this criterion are “Do the learners mainly use the language for a communicative purpose, not just learn the language (e.g., forms and functions) intentionally?” (Q1.1) and “Is the learner primarily concerned with producing and/or comprehending the messages (semantic and pragmatic meaning), not with focusing on linguistic form” (Q1.2) (Jaruteerapan, 2020, p. 71). Therefore, the activities that mainly focus on linguistic forms or language in isolation were considered non-tasks (i.e., exercises). Below are the examples of a non-task in Figure 7 (Çimen et al., 2019b, p. 115) and a task in Figure 8 (Bulut et al., 2019, p. 21), which differ in terms of their main focus. When examining the activities below, it can be observed that both of the activities are input-based and require learners to perform tasks receptively. On the other hand, they differ in terms of how learners are expected to deal with language.

Figure 7

Theme 9 Activity 5b in Count me in for Grade 10

5 B

Below are some phrases related to imaginary situations. Listen to the dialogue and tick the ones you hear. Tapescript 9.1

- ☐ I wish
- ☐ I suppose
- ☐ As if
- ☐ If only
- ☐ If I were you
- ☐ I would

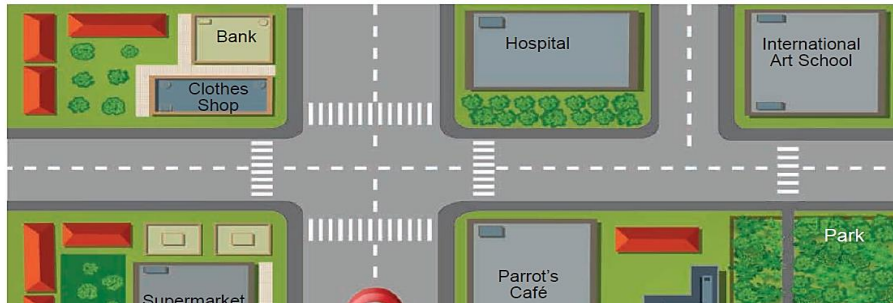
Note. From *Count me in 10 (Student's Book)* (p. 115), by F. Çimen, B., Taşkıran Tiğın, A. Çokçalışkan, N. Özyıldırım, M. Özdemir, E. Demir, & G. Fındıkçı, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

In the first activity (Figure 7), learners do not necessarily need to process the meanings of the words for a communicative purpose (Q1.1). On the contrary, all they have to do is to listen to the recording for specific phrases and put a tick next to the ones they hear. As they can easily complete the activity without attending to either semantic or pragmatic meaning (Q1.2), such activities were counted as non-task.

Figure 8

Theme 1 Activity 6 in Teenwise for Grade 9

6  Listen to the dialogue. Follow the directions and draw a line from Tom to Eva's house.



Note. From *Teenwise 9 (Student's Book)* (p. 21), by E. Bulut, F. Baydar Ertopçu, S. Umur Özadali, & S. Şentürk, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

The second activity above (Figure 8) is a typical listen-and-do activity, which requires learners to act on what they hear by using their own (non)linguistic resources to comprehend messages and paying attention to both semantic and pragmatic meanings of words (Q1.1) so that they can close the gap and complete the activity in question. In this way, they use the target language for a communicative purpose instead of primarily focusing on form or function (Q1.2). As the activity meets the other criteria of taskness as well, it was categorised as a *task*.

Criterion 2

The second criterion requires some kind of *gap* (Ellis, 2009; Ellis & Shintani, 2014). The two questions asked for this criterion are (Q2.1) “Does the activity have a gap that requires a transfer of a given message from one person to another, or one form to another (written to oral) or one place to another (texts to table) in order to close it?” and (Q2.2) “Is the gap closed as a result of learners’ participation in communication (i.e., as either as a speaker/writer or listener/reader)?” (Jaruteerapan, 2020, p. 71).

Though some activities have a resemblance to tasks and meet some of the criteria for taskness, they were counted as non-task (i.e., exercise) as there is no genuine gap to be closed. For instance, describing one’s appearance when it is obvious to both parties (Peng & Pyper, 2021) or ordering the pictures of a story when all the task-takers already know the correct order (Lambert, 2018) were considered non-tasks in the present study. Below are the examples of a non-task in Figure 9 (Bulut et al., 2019, p. 28) and a task in Figure 10 (Akdağ et al., 2019, p. 84) that differ in terms of criterion 2.

Figure 9

Theme 2 Activity 2 in Teenwise for Grade 9

2 Work in pairs. Look at the picture below. Ask and answer questions as in the example.



Note. From *Teenwise 9 (Student's Book)* (p. 28), by E. Bulut, F. Baydar Ertopçu, S. Umur Özadali, & S. Şentürk, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

Although there seems to be a transfer of information in the first activity (Figure 9), it lacks a gap to be filled as learners are required to ask and answer questions about a picture both of them can see (Q2.1). Thus, there is not a genuine gap that needs to be closed as a result of learners' participation in communication (Q2.2). As the activity does not meet criterion 2, it was marked as *non-task*.

Figure 10

Theme 6 Activity 7 in Silver Lining for Grade 11

7 Work in groups. Go to your own page to look at the big picture. Then describe it to the other groups.

Group A: Go to page 134.
Group B: Go to page 135.

Group C: Go to page 136.

Note. From *Silver Lining 11 (Student's Book)* (p. 84), by E. Akdağ, F. Baydar Ertopçu, K. Uyanık Bektaş, S. Umur Özadali, & T. Kaya, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

The second activity (Figure 10) is performed after students read three stories purposefully left incomplete. Upon reading the first part of the stories in the previous activity, students are expected to read the rest of the stories on different pages that their group members do not see and narrate the end of the stories to one another. In this case, there is an information gap to be filled (the end of the story) (Q2.2), and learners find out the rest of the stories as a result of the transfer of a message from one person to another (Q1.2). As the activity also satisfies the other criteria for tasks, it was marked as a *task*.

Criterion 3

The third criterion is that learners rely mainly on their own (non) linguistic resources (Ellis, 2009; Ellis & Shintani, 2014). It means that a task does not present, teach or dictate a specific linguistic form to be utilised during task performance as learners are expected to employ their own (non)linguistic resources for decoding or encoding messages. Two questions utilised for this criterion were (Q3.1) “Are the learners not “taught... the language they will need to perform the activity?” and (Q3.2) “Does the activity allow the learners to use the language they already knew to complete the activity?” (Jaruteerapan, 2020, p. 71).

With this criterion in mind, the activities that explicitly present linguistic forms expected to be utilised or sample language to get help from were counted as *non-task*. These activities generally provide linguistic information in the task instructions (e.g., “use...”) or next to the task content, generally accompanied with examples marked with “e.g.”. Such information is mainly used to direct the learners’ attention to specific language forms by preventing them from primarily drawing on their own (non)linguistic resources while performing a language activity. On the other hand, the activities with such expressions were not immediately considered as

non-task. If the example or model was not there to influence the use of specific linguistic forms, but to provide learners with a sample in general for meaning, then the activities with such examples or models were considered as tasks if they also met the other three criteria.

The examples of a *non-task* in Figure 11 (Bulut et al., 2019, p. 81) and a *task* in Figure 12 (Akdağ et al., 2019, p. 96) are given, which possess different characteristics. In both activities, learners are provided with a model for writing. While determining whether an activity is a task or a non-task concerning criterion 3, it was paid attention to the function of the example given.

Figure 11

Theme 6 Activity 6a in Teenwise for Grade 9

- 6** a. Work in pairs. Prepare a similar dialogue to buy a plane ticket to a foreign country as in the example. Then act it out.

Student A : You are the passenger. Answer the travel agent's questions to buy a ticket.

Student B : You are the travel agent. Ask the passenger the name, the destination, the date and give the price.

A : Good morning. I'd like to buy a ticket to *(your destination)*, please.
 B : Sure. What is your name, please?
 A : My name is *(your name & surname)*.
 B : OK. Mr./Ms. *(the passenger's surname)*. Which date would you like to travel?
 A : *(the travel date)*.
 B : OK. Let me check. One-way or round trip?
 A : *(One-way / Round)* please. How much is the ticket?
 B : From *(your city)* to *(destination)*, *(one-way / round)* trip, *(economy / business)* class...
 It's *(the price)*. You are leaving on *(day & time)* and arriving *(day & time)*. Here is your ticket.
 A : OK. Thank you. Goodbye.

Note. From *Teenwise 9 (Student's Book)* (p. 81), by E. Bulut, F. Baydar Ertopçu, S. Umur Özadali, & S. Şentürk, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

In activity 6a above (Figure 11), it is seen that learners are expected to write a dialogue similar to the one given as an example. In the example, there are some formulaic expressions learners can easily use when completing their dialogues without using their own linguistic resources. In this case, learners are somehow taught the language and expected to recreate the example with small changes (Q3.1), thus not being allowed to use the language they already know to complete the

activity (Q3.2). Therefore, the contents of the model function as explicit language as in *non-tasks*.

Figure 12

Theme 7 Activities 7a-b in Silver Lining for Grade 11



Note. From *Silver Lining 11 (Student's Book)* (p. 96), by E. Akdağ, F. Baydar Ertopçu, K. Uyanık Bektaş, S. Umur Özadalı, & T. Kaya, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

As for the activity 7a (Figure 12), it was considered a task due to its focus on meaning, the existence of a gap, learners' reliance on their own (non) linguistic resources and a specified outcome that determines the completion of the task (composing an email text). There is no teaching of any language to learners (Q3.1) and they are given the freedom to use the language they already know to complete the activity (Q3.2). Therefore, such activities were counted as *task* in this study.

Criterion 4

The last criterion dictates that “there is a clearly defined communicative outcome” (Ellis, 2009, p. 223; Ellis & Shintani, 2014, p. 135). That is, the use of language should function as a means to a communicative outcome rather than its correct practice being an end itself during task performance (Ellis, 2009, 2011; Ellis & Shintani, 2014). Two questions asked for this criterion were (Q4.1) “Are the learners primarily concerned with achieving the goal stipulated by the activity, rather than using language forms correctly?” and (Q4.2) “Is there an outcome that results from completing an activity that works towards a communicative goal, other than the display of linguistic knowledge?” (Jaruteerapan, 2020, p. 71).

For this criterion, it would be useful not to forget that even if there is an outcome, it should concern *text creation*, not *text manipulation*. While the first one concerns how information in L2 is manipulated, the latter appertains to how L2 itself is manipulated.

The researcher looked for a well-defined outcome (e.g., drawn map, reached conclusion, a list, a decision made and a winner) to signal the termination of a task performance no matter what the type of a task was in the textbook. One challenge experienced during data analysis was to find out the outcome of open tasks with divergent goals. Opinion gap tasks fall under this category, which is generally performed in the form of a free flowing conversation without a clear, observable outcome (Cobb, 2010). Therefore, it was paid attention to mark the opinion gap activities as a task when they are more structured (i.e., a procedure to follow) and have a specific outcome that shows the task is accomplished as Cobb (2010) did. On the condition that an opinion gap task necessitates moving beyond the mere presentation of ideas through such processes as the justification, evaluation of the

ideas of others, comparison of positions and ranking of points made in a structured way (Willis & Willis, 2007), then it was accepted as a task.

The examples of a *non-task* activity in Figure 13 (Çimen et al., 2019b, p. 127) and a *task* in Figure 14 (Çimen et al., 2019a, p. 90) are provided below.

Figure 13

Theme 10 Activity 5c in Count me in for Grade 10

5 C

Work in pairs. Choose one of the role cards below and act out your dialogues.

Shop Assistant

- Ask whether he/she needs help.
- Say the item is large, ask what size s/he wants.
- Say that you've found the XL.
- Say the changing rooms are over there.
- Ask how the item fits.
- Tell the price.

Customer

- Ask for the size of the item.
- Say you want an XL.
- Thank him/her and ask where you can try it on.
- Thank him/her.
- Say it suits you and ask its price.
- Say you'll buy it and thank.

Note. From *Count me in 10 (Student's Book)* (p. 127), by F. Çimen, B., Taşkıran Tiğın, A. Çokçalışkan, N. Özyıldırım, M. Özdemir, E. Demir, & G. Fındıkçı, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

Activity 5c (Figure 13) is a cue-card activity that allows learners to interact orally by partially drawing on their (non)linguistic resources to close the gap. The activity seems to provide situational authenticity as learners can engage in such conversations outside the school context in real life. However, it is seen that it directs learners to focus on the linguistic forms that serve the intended language functions rather than providing them with an opportunity to focus on the task goal (Q4.1). In addition, it is hard to differentiate the outcome of the activity from the language that is produced during the performance, which makes the performance of the activity the only outcome (Q4.2). Although the activity has a strong resemblance to tasks, it was treated as *non-task*.

Figure 14*Theme 9 Activity 5e in Count me in for Grade 12*

E. The text is a combination of extracts from websites, brochures and magazines on high-tech products. Read it and list the most necessary gadgets for each of the people given. Consider their functions while categorizing. More than one option is possible.

Smartphones can run applications on an operating system. They are cordless telephone sets and equipped with computer-enabled features that are not previously used by their predecessors. They can be used for both communication and many computer related operations.

Tablet PCs are generally called tablets for short. They are mobile and have touchscreen panels. They also have mobile operating systems that are capable of doing almost the same things as PCs. They lack some I/O ports, though. They are somewhere between smart phones and laptop computers.

Smart watches are portable devices that you can wear on your wrist, just like traditional school wrist watches. They have touch screens and operating systems that allow you to run programs for many purposes such as calculations, digital time telling, translations, and game playing.

Quadcopters are like helicopters but they are much smaller and have four propellers attached to rotors. They are used across vast areas for transportation, surveillance etc. purposes. The terms drone and quadcopter are mostly used interchangeably. However, they are different devices though they have some certain similarities.

Interactive touch-screen projector is a projector that turns any surface into an interactive touch screen. This means you can annotate a presentation, play some games in huge dimensions, or beam a recipe on to your chopping board and scroll through it with doughty finger.

Smart cameras are security cameras that automatically identify human beings and even recognize their faces alerting you via your phone. They also have infrared night vision, live warning speakers and multi-axis rotation systems.

Robot vacuum is better for pet hair, has an application that allows you to control it remotely. It can access to no-go areas.

University professor	Movie Director	Housekeeper	Mountaineer

Note. From *Count me in 12 (Student's Book)* (p. 90), by F. Çimen, B. Taşkiran Tiğin, A. Çokçalışkan, N. Özyıldırım, & M. Özdemir, (2019), Milli Eğitim Bakanlığı. Copyright 2019 Milli Eğitim Bakanlığı.

On the other hand, activity 5e in Figure 14 is considered to satisfy the *outcome* criterion. Learners are expected to focus on achieving the task goal instead of correctly using language forms (Q4.1). The completion of the activity also serves a communicative goal (e.g., a list prepared) rather than the display of linguistic knowledge (Q4.2). Since the activity also satisfies the other task criteria, it was categorised as *task*.

Based on the four criteria explained above, the activities were classified into tasks and non-tasks (i.e., exercises) based on the framework by Ellis (2009), Ellis and

Shintani (2014) and Jaruteerapan (2020). Since tasks were the focal point in the second and third stages of the present study, non-tasks were excluded from the following stages of analysis in the present study.

Analysis of Task Types

This section is devoted to describing how the tasks in the textbooks were analysed based on the task typology proposed by Ellis (2003, 2012, 2020). It includes five dual categories named (1) *input-based* versus *output-based* tasks, (2) *real-world* (*real-life*) versus *pedagogic* tasks, (3) *focused* versus *unfocused* tasks, (4) *here-and-now* versus *there-and-then* tasks and (5) *closed* versus *open* tasks. Sample Analysis for the Identification of Task Types randomly chosen from each textbook examined in the present study is given below (Figure 15).

Figure 15

Sample Analysis for the Identification of Task Types

No	G	U	T	FO	UN	RW	PE	IB	OB	*INT	HN	TT	CL	OP
6	9	2	Read the text about William's room and complete the sketch of it.	1	0	0	1	1	0	0	1	0	1	0
39	10	1	Work in pairs. Discuss the ways of integrating the everyday habits in the reading in 5 B into your everyday lives and decide how applicable they are. Then, draw your percentage line on the diagram as in 5 D	1	0	0	1	0	0	1	1	0	0	1
83	11	3	Listen to the dialogue between Colt and his father, Steve. Put the pictures in the correct order.	1	0	0	1	1	0	0	1	0	1	0
119	12	3	Imagine that you are members of the university human rights club. The president Mrs. Merits requests you to create mottos and slogans about human rights in addition to the ones below, already existing in the club's official website. Work in groups to create new mottos and slogans.	0	1	1	0	0	0	1	0	1	0	1

Note. G = grade, U = unit, T = task, FO = focused, UN = unfocused, RW = real-world, PE = pedagogic, IB = input-based, OB = output-based, *INT = interactive, HN = here-and-now, TT = there-and-then, CL = closed, OP = open.

As seen in Figure 15, each task was analysed based on five dual categories in an Excel spreadsheet with relevant information regarding the number of the tasks (No) as assigned by the researcher, grade (G), unit (U), task (T) as well as task types, which are (1) *input-based* (IB) versus *output-based* tasks (OB) and *integrative* tasks (INT), (2) *real-world* (RW) versus *pedagogic* tasks (PE), (3) *focused* (FO) versus *unfocused* tasks (UN), (4) *here-and-now* (HN) versus *there-and-then* tasks (TT) and (5) *closed* (CL) versus *open* tasks (OP).

Integrative task (INT) type was added to the checklist during the task implementation process as some tasks were observed to necessitate the use of skills from both input-based and output-based tasks. This type was marked with an asterisk (*) symbol to indicate that this category emerged during the analysis process.

To mark the checklist based on the task types, a binary coding system of 0 and 1 was utilised, 1 representing the presence of a task type characteristic while 0 standing for its absence.

When coding was done, the frequencies and percentages of each task type in the textbooks were calculated. The following sections explicate how coding was performed regarding each task type.

Focused versus Unfocused Tasks

The first category distinguishes tasks depending on whether they orient learners to using specific linguistic features while not explicitly presenting these features (*focused*) or allowing them to use language communicatively in general (*unfocused*). Two criteria were paid attention to when making decisions. First, when a task made a pre-determined linguistic feature *very useful* or a *must* to be completed or when this linguistic form was highly likely to occur during task implementation naturally, such tasks were considered *focused* (Ellis, 2003). Second, it was necessary

to find out if this linguistic form was targeted within the relevant unit. With this aim in mind, the researcher examined the contents, objectives and outcomes of each unit specified when making relevant decisions as to whether the tasks are FO or UN.

While the example of a FO task can be found in Figure 16 (Bulut et al., 2019, p. 29), an UN task can be found in Figure 17 (Akdağ et al., 2019, p. 93).

Figure 16

Theme 2 Activity 3 in Teenwise for Grade 9

3 Read the text about William's room and complete the sketch of it.

MY LITTLE WORLD

This is my room. It is my favourite place and my only private space. There is a dark blue carpet on the floor. There is a window near my bed. I can see the beautiful garden from my bed. There are three posters of comic book characters on the walls.

There is a bedside table next to my bed. My books and alarm clock are always on it. On the left side of the bed, there are two shelves on the wall for my model planes. My wooden box under the bed is the most important thing in my room. My precious things are in it - all my diary. They are very important to me.



Note. From *Teenwise 9 (Student's Book)* (p. 29), by E. Bulut, F. Baydar Ertopçu, S. Umur Özadali, & S. Şentürk, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

In the first task (Figure 16), learners are expected to read a text and complete a sketch of it. The task counts as *focused* since the textbook already targets prepositions of place within the scope of the theme that task-takers need to deal with to accomplish the task.

Figure 17

Theme 7 Activity 7 in Silver Lining for Grade 11

7 Search the Net for a myth. Follow the given steps to prepare a presentation.

- a. Choose a myth.
- b. Read and take notes.
- c. Find visual materials.
- d. Prepare your presentation using different Web 2.0 tools.

Note. From *Silver Lining 11 (Student's Book)* (p. 93), by E. Akdağ, F. Baydar Ertopçu, K. Uyanık Bektaş, S. Umur Özadalı, & T. Kaya, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

In the second task (Figure 17), learners are required to search for a myth online and prepare a presentation. As learners are not directed to use any specific linguistic feature to complete the task, it was marked as *unfocused*. Although past tense would be practical for learners to report past events, for instance, it was not one of the pre-determined linguistic features in Theme 7.

Real-world versus Pedagogic Tasks

The second category is related to whether a task has situational authenticity and corresponds to real-life events (*real-world*) or only has interactional authenticity but still functions to facilitate natural language (*pedagogic*). For the real-world task category, the researcher considered if a similar task can be practised outside the school context for real-life purposes (e.g., booking a hotel room) or not. A real-world task (Figure 18) from the textbook for grade 12 (Çimen et al., 2019a, p. 70) and a pedagogic task (Figure 19) from the textbook for grade 9 (Bulut et al., 2019, p. 20) from the textbook from grade 9 are given below.

Figure 18

Theme 7 Activity 2b in *Count me in* for Grade 12

B. Imagine that you are a reporter and you have interviewed Oumar Hoidini, an earthquake survivor, for a story in your newspaper and below is what he told you about his experience. Paraphrase the story before submitting it to your editor.

The earthquake hit the area while most were asleep, just before the dawn, at a September night. There was no electricity, no food, no water, no mobile phone signals. All I had under the debris was a small space and oxygen.

It was far too late when I noticed the truth. I mean, the earthquake. The building had collapsed. Even though I was in a terrible situation, I tried to remain calm and patient. I never lost hope. I don't know why. Perhaps, it was because I believed that some people, rescue teams or one of the beloved ones could get me out of the building. I just sat down and waited. I had time for reflection about what my life had been like. I also thought about how I would appreciate each moment of life if I survived. Suddenly, I heard some

Note. From *Count me in 12 (Student's Book)* (p. 70), by F. Çimen, B., Taşkiran Tiğin,

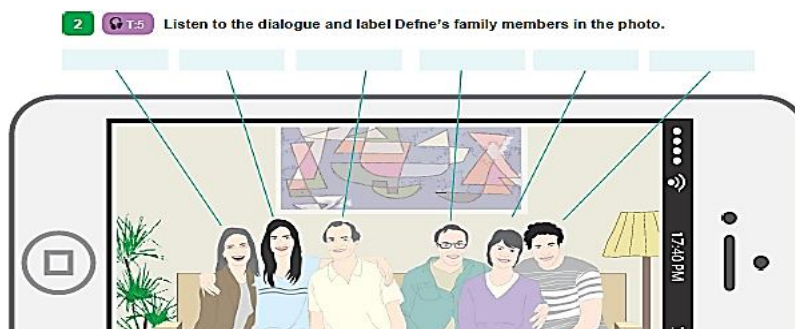
A. Çokçalışkan, N. Özyıldırım, & M. Özdemir, 2019, Milli Eğitim Bakanlığı.

Copyright 2019 by Milli Eğitim Bakanlığı.

For the first task (Figure 18), task-takers are instructed to read the experiences of an interviewee and report them by paraphrasing his utterances. Such a task both facilitates natural language that occurs in L2 settings (interactional authenticity) and corresponds to real-life events (situational authenticity) as a reporter can perform the same task on a daily basis. Therefore, the activity was marked as *real-world* task.

Figure 19

Theme 1 Activity 2 in *Teenwise* for Grade 9



Note. From *Teenwise 9 (Student's Book)* (p. 20), by E. Bulut, F. Baydar Ertopçu, S.

Umur Özadali, & S. Şentürk, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli

Eğitim Bakanlığı.

The second activity (Figure 19), on the other hand, engages learners in listening to a recording that would occur in a non-pedagogic conversation (interactional authenticity); yet, it fails to have situational authenticity as it is not likely that a person writes down the names of the people in a photo while listening to a description of them without a plausible reason to do so outside the school.

Input-based versus Output-based Tasks

The third category deals with whether a task necessitates only decoding a message and requires no or minimal verbal production (input-based) or necessitates oral or written language production (output-based). During task analysis, minimal verbal production was operationalised as single words (e.g., “Yes”, “Sure”) and phrasal utterances (e.g., “No way!”) and short clauses (“I disagree”, “to see a doctor”) upon receiving an expert opinion. Then, the tasks that required minimal verbal production or no verbal production at all were counted as *input-based*. On the other hand, the tasks that require learners to speak or write (except for minimal response) were labelled as *output-based*. Finally, when a task engages learners in at least two skills from two different categories (input-based and output-based) at to accomplish the task, then it was coded as *integrative* (INT). The examples of *input-based* (Çimen et al., 2019b, p. 119), *output-based* (Çimen et al., 2019a, p. 82) and *integrative* tasks (Çimen et al., 2019b, p. 19) are provided in Figure 20, Figure 21 and Figure 22, respectively.

Figure 20

Theme 10 Activity 1c in Count me in for Grade 10

1 C

There are four dialogues between a cargo delivery guy and an officer, in each of which one person is described. Listen to them and find out which person is related to which dialogue. Then, put the pictures into the correct order. Tapescript 10.1

Note. From *Count me in 10 (Student's Book)* (p. 119), by F. Çimen, B., Taşkiran Tiğın, A. Çokçalışkan, N. Özyıldırım, M. Özdemir, E. Demir, & G. Fındıkçı, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

The first activity (Figure 20) is an *input-based* task as it simply requires learners to listen to the descriptions of the people and find out to whom the cargo delivery staff needs to deliver stuff. The only output is the numbers learners write next to the images of the people.

Figure 21

Theme 8 Activity 6c in Count me in for Grade 12

C. Work in pairs. Imagine that one of you is a person who wants to make a complaint about an environmental problem and the other is a local authority who will offer solutions. Make a dialogue and share it with your friends.

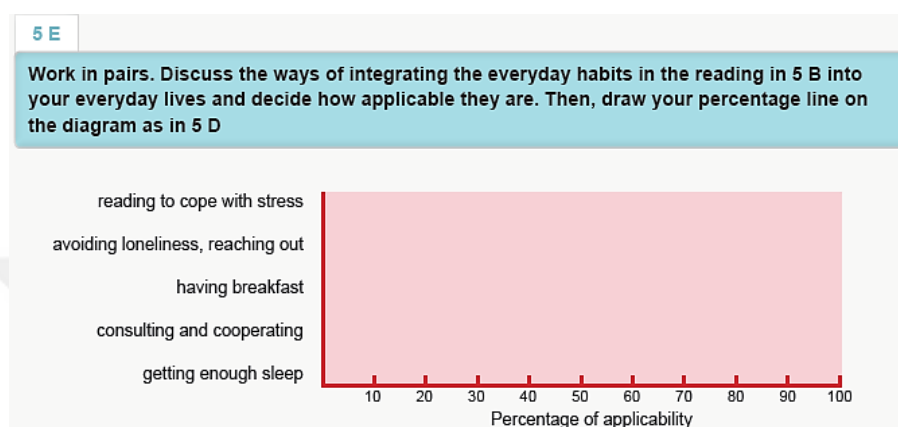
Note. From *Count me in 12 (Student's Book)* (p. 82), by F. Çimen, B., Taşkiran Tiğın, A. Çokçalışkan, N. Özyıldırım, & M. Özdemir, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

The activity above (Figure 21) involves using a productive skill (i.e., speaking) to role-play a dialogue, in which one task taker assumes the role of a

person who wants to make a complaint about an environmental problem and the other a local authority to offer solutions. The task, therefore, qualifies as *output-based* task.

Figure 22

Theme 1 Activity 5e in Count me in for Grade 10



Note. From *Count me in 10 (Student's Book)* (p. 19), by F. Çimen, B., Taşkiran Tiğın, A. Çokçalışkan, N. Özyıldırım, M. Özdemir, E. Demir, & G. Fındıkçı, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

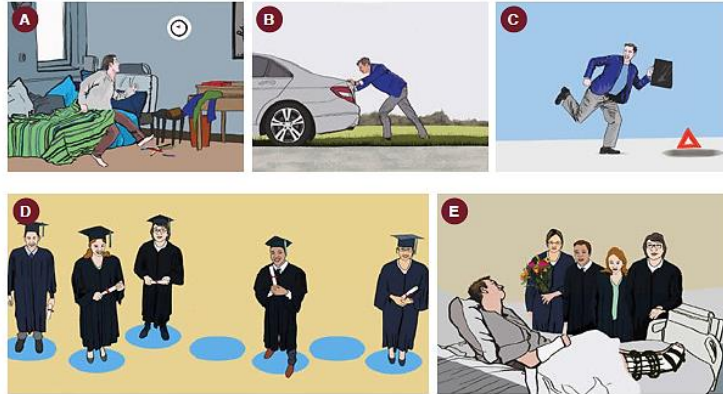
The third activity (Figure 22), on the other hand, is a good example of an *integrative* task as it engages learners in reading, listening (*input-based*) and speaking (*output-based*) as they need first to read a text and then, share information/opinion in a discussion to complete the diagram collaboratively.

Here-and-now versus There-and-then Tasks

Regarding the fourth category, it concerns whether learners are provided with the task-related content in front of them (*here-and-now*) or have no access to it (*there-and-then*). If task-takers can access the task content, then such tasks were categorised as *here-and-now*, while the rest were marked as *there-and-then*. Figure 23 (Akdağ et al., 2019, p. 81) provides an example of a *here-and-now* task, while Figure 24 (Çimen et al., 2019a, p. 88) presents a *there-and-then* task.

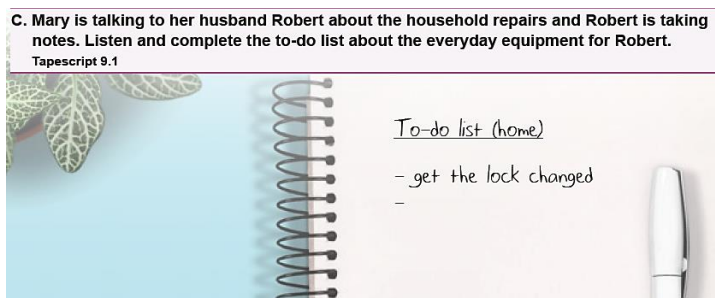
Figure 23*Theme 6 Activity 6 in Silver Lining for Grade 11*

6 Work in groups. Make deductions for each picture and note them down to create a story. Share your story with the other groups.



Note. From *Silver Lining 11 (Student's Book)* (p. 81), by E. Akdağ, F. Baydar Ertopçu, K. Uyanık Bektaş, S. Umur Özadalı, & T. Kaya, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

The first activity (Figure 23) provides learners with a set of pictures, from which they need to make deductions to create a story. Since they are allowed to keep the images to check the details on and off during task implementation, the activity was considered a 'here-and-now' task.

Figure 24*Theme 9 Activity 2c in Count me in for Grade 12*

Note. From *Count me in 12 (Student's Book)* (p. 88), by F. Çimen, B., Taşkiran Tiğın, A. Çokçalışkan, N. Özyıldırım, & M. Özdemir, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

As for the second activity (Figure 24), it is a *there-and-then* task since the task-takers only depend on the recording they can listen to. It means they need to listen to the speaker and note the activities mentioned but cannot see them. Therefore, such activities were taken as a *there-and-then* task within the scope of the present study.

Closed versus Open Tasks

Finally, the last category deals with the number of task outcomes. While some tasks can have a single or limited number of outcomes (*closed*), some may have several outcomes (*open*). An example of a closed task can be found in Figure 25 (Çimen et al., 2019b, p. 68), while an open task can be found in Figure 26 (Çimen et al., 2019b, p. 105).

Figure 25

Theme 5 Activity 6b in Count me in for Grade 10

A receptionist is calling Ethan to confirm a table reservation. Look at the notes taken by the receptionist below. Listen to the dialogue and correct the notes that are taken incorrectly. Tapescript 5.4



Note. From *Count me in 10 (Student's Book)* (p. 68), by F. Çimen, B., Taşkiran Tiğın, A. Çokçalışkan, N. Özyıldırım, M. Özdemir, E. Demir, & G. Fındıkçı, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

In the first activity (Figure 25), the task takers are instructed to listen to the recording and correct the wrong information given. There is only one acceptable word, number etc. they need to find out to correct the information on the note. Hence, it was marked as a *closed* task.

Figure 26

Theme 8 Activity 6e in Count me in for Grade 10

6 E

Below is the template of a digital collaborative story. Its first sentence has been given for you. Each time before a group, pair or student adds a sentence, discuss it and make changes if necessary to create a collaborative story which shows the advantages of technology. Then, arrange it online.

First line	They realized that they were lost.
Add sentence	
Add sentence	
Add sentence	
Add sentence	
Add sentence	
Add sentence	
Last line	

Note. From *Count me in 10 (Student's Book)* (p. 105), by F. Çimen, B., Taşkıran Tiğın, A. Çokçalışkan, N. Özyıldırım, M. Özdemir, E. Demir, & G. Fındıkçı, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

The second activity (Figure 26) is an *open* one since task-takers are allowed to write their responses by including whatever information they would like as long as they are relevant to the previous sentences written by other groups.

Analysis of Task Complexity Dimensions

Checklist 3, focused on the cognitive demands of the output-based tasks on language learners. It was constructed drawing on two frameworks, *Triadic Componential Framework* by Robinson (2001a, 2001b, 2003, 2007, 2008, 2011, 2015) and *Task Difficulty Scheme* by Skehan and Foster (2001). Robinson's TCF was used as a base in the first place to create the checklist, and the one by Skehan and Foster was utilised to explicate one dimension from Robinson's framework.

The final checklist (Checklist 3) for task complexity dimensions is provided below (Figure 27).

Figure 27

Sample Analysis for the Exploration of Task Complexity Dimensions

No	G	U	T	Resource-directing Dimensions										Resource-dispersing Dimensions										SSARC Steps																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
				(+)		TT		(+)		SR		(+)		IR		PT		(-)		ST		(-)		TS		(-)		PK		Step 1	Step 2	Step 3																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
				(+)	(+)	*TIM	*STA	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)

Note. G = grade, U = unit, T = task, TT = there-and-then, *SPA = space, *TIM = time, SR = spatial reasoning, *STA = static,

*DYN = dynamic, CR = causal reasoning, IR = intentional reasoning, PR = perspective-taking, ST = single task, *COS = conceptual

support, *MUL = multitasking, TS = task structure, *CON = context, *STE = steps, *SS = single step, PK = prior knowledge,

*TOP = topic, *DIG = discourse genre, *CCD = cognitive / contextual demands, *RDir = resource-directing, *RDis = resource-dispersing.

As demonstrated in Figure 27, Checklist 3 consists of two categories (i.e., *resource directing* and *resource dispersing* dimensions) and nine sub-categories for task complexity, the first five belonging to *resource directing* dimensions, while the other four belonging to *resource dispersing* dimensions. It also has a third section for task complexity levels of tasks based on SSARC model.

Before the introduction of the categories and sub-categories in Checklist 3, it must be highlighted that the complex version of each task complexity was represented there. That is, while the *resource directing* dimensions are marked with (+), *resource dispersing* dimensions were marked with (-) since the presence of the first group of dimensions increases task complexity, the absence of the latter does the same effect.

The first task complexity category in Checklist 3, *resource directing* dimensions, consists of five task conditions: (1) +TT, (2) +SR, (3) +CR, (4) +IR and (5) +PT. The second task complexity category, *resource dispersing* dimensions, includes four task conditions, which are (1) -ST, (2) -TS, (3) -SS and (4) -PK.

The last section of the Checklist 3 demonstrates the complexity levels of tasks as defined in SSARC (“stabilize”, “simplify”, “automatize”, “restructure”, “complexify”) model (Robinson, 2015, p. 12), which are Step 1 (-Rdir -RDis), Step 2 (-RDir +RDis), Step 3 (+RDir +RDis) and a new, intermediary step generated between Step 2 and 3 within the scope of the present study (+RDir -RDis).

To mark Checklist 3 based on the cognitive complexity of tasks, a binary coding system of 0 and 1 was utilised, 1 representing the presence of a task complexity dimension while 0 standing for its absence. Besides, depending on along which task complexity dimension(s) a task was found to be made complex, the task complexity steps were marked with 1 and 0. If a step was marked with a 1, it

indicates that a task was detected to be in this level of complexity. Otherwise, the step was marked with 0. When coding was done, the numbers of all the 1s and 0s were calculated to find the frequencies and percentages of each task complexity dimension and task complexity level.

The following sections illustrate how tasks were analysed in the current study based on the *resource-directing* and *resource-dispersing* dimensions.

Resource-directing Dimensions

The resource-directing dimensions examined in the present study include *there-and-then* (TT), *spatial reasoning* (SR), *causal reasoning* (CR), *intentional reasoning* (IR) and *perspective-taking* (PT). They are explained in the following sections.

There-and-then (TT). TT condition is concerned with the “cognitive operations, conversational abilities, and linguistic resources” (Robinson, 1995, p. 102) needed for making displaced reference to time and space and, thus, directing attention to form-function mappings in the L2 during task implementation (Robinson, 2014). For this condition, tasks were analysed to determine if they necessitated displaced references to (1) *space* (SPA) and (2) *time* (TIM), increasing cognitive/conceptual complexity. It must be noted that when making decisions regarding The TT dimension, the researcher paid attention to whether a displaced reference to events was necessary (+TT) or not (-TT) regardless of the presence or absence of a task input. The analysis of a task based on the TT dimension is given in Figure 28 (Akdağ et al., 2019, p. 42).

Figure 28*Theme 3 Activity 9 in Silver Lining for Grade 11*

- 9 Look at Adam's picture and read the beginning of his story. Then, choose your favourite ending from picture A or B and complete the story.



Adam was a poor little boy. He used to live in a small house with his big family. He couldn't attend school regularly because he had to work and support his family ...

.....

.....

.....

.....

.....




Note. From *Silver Lining 11 (Student's Book)* (p. 42), by E. Akdağ, F. Baydar Ertopçu, K. Uyanık Bektaş, S. Umur Özadalı, & T. Kaya, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

In the task (Figure 28), learners are instructed to complete a story based on the (non)linguistic input provided. Since the story they are to write is dislocated in time and space, it was taken as a +TT task. Please note that although they can access the visual and non-visual task input during the task implementation process, this was not considered when coding it in terms of TT condition, as it concerns contextual support, which is considered within the scope of *single task* in resource-dispersing dimensions.

Spatial Reasoning (SR). ST dimension deals with if a task induces the employment of linguistic forms to describe motion events and/or location. For this reasoning type, it was paid attention to whether learners are expected to only transmit information through the description of a picture/map given (-SR) or make decisions

Figure 30*Theme 8 Activity 3f in Count me in for Grade 10*

3 F

Read the following statements related to using the Internet and social media and discuss to find out the problems.

1. A friend of mine tagged me in a picture without permission.
2. While I was reading comments on a forum page, I saw some offensive and foul language.



netiquette noun net-i-quette \ 'ne-ti-ket , -ket \

rules about the proper and polite way to communicate with other people when you are using the Internet.

Note. From *Count me in 10 (Student's Book)* (p. 102), by F. Çimen, B., Taşkiran Tiğın, A. Çokçalışkan, N. Özyıldırım, M. Özdemir, E. Demir, & G. Fındıkçı, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

The task (Figure 30) engages learners in causal reasoning by thinking about causes and effects of problems. Task-takers are likely to use specific expressions like logical subordinators (e.g., *so*, *because*) to justify their interpretations regarding why one event leads to another (Robinson, 2005). Therefore, it was considered to pose causal reasoning demands (+CR).

Intentional Reasoning (IR). IR dimension is concerned with whether tasks require learners to reason about the motives, intentions, beliefs or desires of people as well as the relationships between these (Robinson, 2007, 2008, 2011; Robinson & Gilabert, 2007) by using relevant linguistic features to explain why they do what they do and what consequences follow (+IR) or not (-IR). The analysis of a task based on the IR dimension is provided in Figure 31 (Akdağ et al., 2019, p. 50).

Figure 31*Theme 4 Activity 2 in Silver Lining for Grade 11***2 Read and talk about the childhood dreams of the celebrities below. Then, match them with their photos.**

- ☐ 1. "Anyone who has spent any time in space will love it for the rest of their lives. I achieved my childhood dream of the sky."
- ☐ 2. "As I always wanted to be a famous author, I began writing and selling monster stories at an early age. I also wrote stories about a mythical kingdom of my pet turtles in toy castles."
- ☐ 3. "At a young age, I told my parents about my decision to study medicine. And since then I've believed what I did in the past is not enough, there is also the future."
- ☐ 4. "After I visited the İstanbul Archaeological Museum at the age of seven, I put all my mother's jewellery in a drawer. Then, I went outside to play 'museum game' with my friends. As you can guess, it didn't last long."



Note. From Silver Lining 11 (Student's Book) (p. 50), by E. Akdağ, F. Baydar Ertopçu, K. Uyanık Bektaş, S. Umur Özadalı, & T. Kaya, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

To accomplish the task goals (Figure 31), learners need to consider these people's motives, intentions or desires and match their childhood dreams with their current jobs, therefore, employing intentional reasoning. For this reason, the task was considered to pose intentional reasoning demands on task-takers (+IR).

Perspective-taking (PT). PT dimension deals with whether task-takers are expected to assume only one first-person perspective on an event (-PT) or several second and third-person perspectives to be able to accomplish a task (+PT) (Robinson, 2007, 2008; Robinson & Gilabert, 2007).

It is important to note that although role plays involve imaginary situations in which task-takers assume the identities of the characters they act out, these tasks were not considered to necessitate perspective-taking since resource-directing

dimensions, including PT, are predicted to direct attention to certain linguistic forms (second and third-person pronouns in perspective-taking dimension) and learners typically assume one-person perspective for role plays rather than second and third-person perspectives. The analysis of a task based on PT dimension is presented in Figure 32 (Çimen et al., 2019b, p. 44).

Figure 32

Theme 3 Activity 7a in Count me in for Grade 10

7 A WRITING

Read the unfinished story below and write an ending for it. Consider the forest in the picture and remember that Jadav Payeng, the hero of the story, has a fame as a legendary forest man in India. The verbs in box may be of help.

plant make water look after begin live grow come back take care

It was 1965 when heavy floods forced the Payengs to move to the other side of the river, 12 km away from their home. The family, with their children, were suffering too much, and when they moved back, they left their five-year-old Jadav in the care of a court-master for schooling. Jadav Payeng gave up school after class 10, and got back to his village to look after the livestock left by his deceased parents.

When he arrived, he witnessed a terrible sight. The land was eroded and there was a sandbar in place of the green land and forests of his childhood. Snakes and reptiles were lying dead on the desert. This sight shook him. Suddenly, he realised the truth. "Today, it is the snakes and tomorrow it will be us," he thought to himself. That moment, he made the decision to make a change, at least to try. He was only 15.

He was determined. He visited the villagers nearby and asked for their opinions. The villagers gave him seeds and shoots of bamboos.

.....

.....

.....

7 B

Now, your teacher will give you the original story in full. Compare your ending with it and find the most similar one in the class.

Note. From *Count me in 10 (Student's Book)* (p. 44), by F. Çimen, B., Taşkiran Tiğın, A. Çokçalışkan, N. Özyıldırım, M. Özdemir, E. Demir, & G. Fındıkçı, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

The task (Figure 32) requires learners to assume a second-person perspective of an event using different pronominal references. Since they are not narrating the story as if it happened to them, but from somebody else's point of view, it was marked to necessitate +perspective-taking (+PT).

Resource-dispersing Dimensions

The resource-dispersing dimensions examined in the present study include *single task* (ST), *task structure* (TS), *single step* (SS) and *prior knowledge* (PK).

They are explained in the following sections.

Single Task (ST). Based on Robinson's framework (2007, 2008, Robinson & Gilabert, 2007), the tasks that require task-takers to do only one thing concurrently during task implementation are coded as single task (+ST), while the ones that necessitate doing two or more things concurrently (+MUL) or performing an activity when the contextual support is absent (-CON) are considered dual task (-ST) within the present study. A sample analysis of a task based on ST dimension is given in Figure 33 (Akdağ et al., 2019, p. 129).

Figure 33

Theme 10 Activity 5a in Silver Lining for Grade 11

- 5 a. Work in pairs. Look at some of the typical behaviours of a grown-up who feels young at heart. Brainstorm and add as many behaviours as you can to the list.
- getting dressed in colourful clothes
 - taking selfies with friends
 - listening to music with headphones on the street
 - feeling the joy inside

Note. From *Silver Lining 11 (Student's Book)* (p. 129), by E. Akdağ, F. Baydar Ertopçu, K. Uyanık Bektaş, S. Umur Özadalı, & T. Kaya, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

For the task (Figure 33) above, task-takers need to both speak and take notes concurrently, which can be considered an example of multitasking (+MUL), a sub-category of the ST dimension. Since they are expected to deal with two things simultaneously during task performance, it was coded as -ST.

Task Structure (TS). Task structure refers to whether there is a clear structure provided so that task-takers can find out the requirements of a task and

distinguish the task steps (+STE) and contents (+CON) or not (-TS) (Robinson, 2007, 2008; Robinson & Gilabert, 2007). The analysis of a task based on TS dimension is provided in Figure 34 (Bulut et al., 2019, p. 90).

Figure 34

Theme 7 Activities 7a-b in Teenwise for Grade 9

7 a. Choose an Ancient Civilization and prepare a presentation of it.

- Search on the Net; the location, history, characteristics and interesting facts.
- Use visuals.
- Write the presentation.
- Use Web 2.0 tools to prepare it.

THE FIRST STEPS OF CIVILIZATION

Mesopotamian Civilization was the first civilization in the world. They lived in the middle-east region of Asia around 3300 BC - 750 BC. The people lived on farming and hunting at those times ...



b. Make a short presentation on the Ancient Civilization you prepared.

Note. From *Teenwise 9 (Student's Book)* (p. 90), by E. Bulut, F. Baydar Ertopçu, S. Umur Özadali, & S. Şentürk, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

In this example, the task (Figure 34) has a pretty clear structure that enables learners to understand what the expectations are as to the contents of the presentation (+CON) as well as the task steps (+STE) to complete the task successfully. Hence, it was coded +TS within the present study.

Single Step (SS). SS dimension is concerned with whether one step is involved in a task (+SS) or more than one step (-SS) to accomplish task outcomes (Robinson, 2007, 2008; Robinson & Gilabert, 2007). A sample analysis of a task based on ST dimension is presented in Figure 35 (Bulut et al., 2019, pp. 104-105).

Figure 35*Theme 8 Activity 3-4 in Teenwise for Grade 9*

3 Do the quiz below and see how much you know about emergency situations.

1. If you see a large fire, you should
 - a) leave immediately
 - b) try to put it out
 - c) call the fire department and leave at once
2. How quickly can a fire spread through a house?
 - a) 20 minutes
 - b) about 5 minutes
 - c) 45 minutes
3. When there is an earthquake, you mustn't be
 - a) in your kitchen
 - b) under a desk
 - c) against a wall
4. Your emergency kit should have enough food and water for at leastdays.
 - a) 2
 - b) 3
 - c) 5
5. Seatbelts reduce the risk of death by
 - a) 30%
 - b) 50%
 - c) 70%
6. If you see a forest fire, call
 - a) 155
 - b) 114
 - c) 110
7. You should call an ambulance
 - a) when you have a minor cut
 - b) when you have the flu
 - c) when you need immediate help

1 or 2 correct answers: You know nothing about emergency. You must definitely google what emergency is. Learn some tips. They can save your life one day.

3-5 correct answers: You have average knowledge about emergency. You should learn more because you may need it.

6 or 7 correct answers: You are a real expert on emergency. You made an excellent job. You know how to handle emergency situations. The people around you are very lucky. You can be their hero one day.

4 Check your results with your classmates. Who is the emergency freak in your class?

Note. From *Teenwise 9 (Student's Book)* (pp. 104-105), by E. Bulut, F. Baydar Ertopçu, S. Umur Özadali, & S. Şentürk, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

For this task (Figure 35), learners first need to do the quiz, calculate their score and then share their results with their classmates and find out the one with the most correct answers. Therefore, learners need to complete more than one step (-SS) to accomplish the task outcomes.

Prior Knowledge (PK). PK denotes background information provided to learners for the task implementation (Robinson, 2007, 2008, Robinson & Gilabert, 2007). It was operationalised as the (1) familiarity of *topic* (TOP), (2) familiarity of

discourse genre (DIG) as well as (3) familiarity of *cognitive/conceptual demands* (CCD) in the present study.

It must be noted that when coding the tasks that require learners to do a web search for an issue, these tasks were considered to provide them with topic knowledge as learners are given a chance to access information that would familiarise them with the topic in question. A sample analysis of a task based on PK dimension is presented in Figure 36 (Akdağ et al., 2019, p. 24).

Figure 36

Theme 1 Activity 7a-b in Silver Lining for Grade 11 (Akdağ et al., 2019, p. 24)

- 7 a.**  Look at Tallynn's CV on page 23 and write a letter of intent for her following the steps below.
- choose a job
 - where / when she saw the job ad
 - details of previous work experience and responsibilities
 - her personal qualities
 - when she is available for an interview and when she can start
- b.** Swap your letters in pairs. If a candidate applied with that letter, would you call him / her for a job interview? Explain your reasons.

Note. From *Silver Lining 11 (Student's Book)* (p. 24), by E. Akdağ, F. Baydar Ertopçu, K. Uyanık Bektaş, S. Umur Özadalı, & T. Kaya, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

The previous task (Figure 36) was coded +PK since it is predicted that learners will be familiar with the topic given, genre-specific requirements as well as cognitive/conceptual demands. Previously, learners will have been given opportunities to think about business-related issues and vocabulary that would familiarise them with the *topic* (TOP) by the time they are asked to implement the task. Besides, they will have examined a model letter of intention and CV and focused on some genre-specific issues (e.g., contents, sections, style) through some activities, promoting familiarity with the *discourse genre* (DIG). Finally, since learners will have already worked on some activities with the TT condition on the

unit, they are expected to be familiar with this condition and be able to make displaced references in space and time while writing the letter of intent as well. Hence, the task was considered to help learners have some prior knowledge (+PK) for the task to be implemented.

SSARC Steps

The final step of the analysis was to detect the task complexity steps depending on along which task complexity dimension(s) a task was found to be made complex. The complexity level of a task in Figure 37 will be examined based on the SSARC model.

Figure 37

Theme 3 Activity 7a in Count me in for Grade 10

7 A WRITING

Read the unfinished story below and write an ending for it. Consider the forest in the picture and remember that Jadav Payeng, the hero of the story, has a fame as a legendary forest man in India. The verbs in box may be of help.

plant make water look after begin live grow come back take care

It was 1965 when heavy floods forced the Payengs to move to the other side of the river, 12 km away from their home. The family, with their children, were suffering too much, and when they moved back, they left their five-year-old Jadav in the care of a court-master for schooling. Jadav Payeng gave up school after class 10, and got back to his village to look after the livestock left by his deceased parents.

When he arrived, he witnessed a terrible sight. The land was eroded and there was a sandbar in place of the green land and forests of his childhood. Snakes and reptiles were lying dead on the desert. This sight shook him. Suddenly, he realised the truth. "Today, it is the snakes and tomorrow it will be us," he thought to himself. That moment, he made the decision to make a change, at least to try. He was only 15.

He was determined. He visited the villagers nearby and asked for their opinions. The villagers gave him seeds and shoots of bamboos.

.....

.....

.....

.....

7 B

Now, your teacher will give you the original story in full. Compare your ending with it and find the most similar one in the class.

Note. From *Count me in 10 (Student's Book)* (p. 44), by F. Çimen, B., Taşkiran Tiğın, A. Çokçalışkan, N. Özyıldırım, M. Özdemir, E. Demir, & G. Fındıkçı, 2019, Milli Eğitim Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

The task in Figure 37 was made complex along both resource-directing (TT and PT) and resource-dispersing (SS) dimensions. As for the resource-directing dimensions, the task necessitates displaced references to space and time (+TT) and assuming a third-person perspective of an event (+PT) to narrate the story. Regarding the resource-dispersing dimensions, more than one step is involved in the tasks to accomplish the task outcomes (-SS).

Since the task is made complex along both resource-directing and resource-dispersing dimensions (+RDir +RDis), its complexity level corresponds to the third and the most complex step of the SSARC model.

Intercoder Reliability Analysis

While making decisions regarding the activities in the textbooks, the researcher got opinions from experts in the field to ensure the right decisions were made. For this reason, two coders, in addition to the researcher, were involved in each step of the analysis of the tasks in the textbooks on a voluntary basis. The coders were two Turkish researchers with MA degrees in the field of ELT and experience in material development and task design.

The coders were trained by the researcher for the scope of the study and the relevant concepts via Zoom meetings prior to each stage of data analysis. Following the training sessions, each coder analysed the relevant data independently. When the coding sessions were over, they held a discussion session to exchange views on their coding.

After each coding and discussion session, the researcher conducted intercoder analysis using Krippendorff's Alpha. Krippendorff's Alpha is "a reliability coefficient developed to measure the agreement among observers, coders, judges,

coders, or measuring instruments drawing distinctions among typically unstructured phenomena or assign computable values to them” (Krippendorff, 2008, p. 1).

Krippendorff’s Alpha was preferred to calculate the intercoder agreement in the present study for two reasons. Although Cohen’s Kappa or Fleiss Kappa are often used for intercoder reliability analysis, they were not preferred in this study as Cohen’s Kappa is used when there are only two coders, and Fleiss’ Kappa requires nonunique coders (randomly selected different raters from the population) for each observation. However, the same three coders performed coding in each stage of the analysis in the present study for practicality. As Krippendorff’s Alpha can be used for any number of coders and any size of data, it constitutes a better tool to compute intercoder reliability analysis (Hayes & Krippendorff, 2007; Krippendorff, 2011).

When a KALPHA value of .80 or above was achieved after Krippendorff’s Alpha reliability analysis, it was considered adequate to continue coding the rest of the data since this value is accepted as a satisfactory level of agreement for categorical (nominal) data (De Swert, 2012; Landis & Koch, 1977).

Intercoder Reliability Analysis for Q1

To measure intercoder (i.e., interrater) reliability among the coders for the analysis data from Q1, two analyses were run, one for the agreement between the coders on task characteristics and the other for their decisions on the taskness of the activities. To answer the Q1, the researcher and the coders examined the activities within a randomly chosen unit in one of the textbooks using Checklist 1 after their training session.

The unit randomly selected to answer the Q1 consists of 23 activities, and the coders analysed these activities against the checklist generated by Jaruteerapan

(2020) with eight questions explicating four criteria of taskness by Ellis (2009) and Ellis and Shintani (2014).

When all the activities in the textbooks were examined, and decisions were made regarding tasks and non-tasks, the researcher held a discussion session with the other coders. During the session, it was discussed how each activity was coded by each coder, in what ways their coding showed similarities and differences, and their justification for how they preferred to code the activities. With the discussion session, it was intended to see what points of discussion as to the coding of activities would arise and how they would be negotiated to facilitate a consensus among coders, reach a higher level of agreement between them and, therefore, strengthen the reliability of coding. During this session, a few points arose related to each criterion.

As for the criterion 1, “the primary focus should be on meaning” (Ellis, 2009, p. 223; Ellis & Shintani, 2014, p. 135), the coders sometimes had difficulty in determining whether learners are expected to use L2 as a means for a communicative purpose or focus on learning the language (e.g., forms and functions) intentionally. This was especially experienced when coding activities that possess ‘interactional authenticity’ but lack ‘situational authenticity’ in Ellis’s terms (2009, p. 227). Specifically, these activities are less likely to be practised in real life (e.g., telling a story using the images given) but enable learners to practice social actions and comprehend/produce language similar to the one occurring in non-pedagogic conversations (Ellis, 2009). While it is pretty easy to detect the focus on meaning in the activities that bear situational authenticity, it can sometimes be hard to determine whether the focus is on meaning (language as a means) or language itself (language as an end) in the activities that have interactional authenticity only.

The second point discussed was related to criterion 2, “there should be some kind of ‘gap’” (Ellis, 2009, p. 223; Ellis & Shintani, 2014, p. 135). The coders had difficulty determining whether there was a gap to be closed in some activities that are typically considered tasks. One example would be the activity in the textbook for 10th graders: “Write an opinion paragraph on one of your plans. You can jot down some notes here.” (Çimen et al., 2019b, p. 31). Although such activities would be called a task in educational settings in general, the coders decided that there is not a communicative gap to be closed, nor is a transfer of a message from one person to another, one form to another or one place to another.

The last point that arose during the discussion session was related to the criterion 4 (outcome). The coders discussed whether the instructions lead learners to produce a communicative (non-linguistic) outcome. They especially negotiated which opinion gap activities would fulfil the requirements of criterion 4. Specifically, it was difficult to determine whether there was an outcome to result from the completion of an activity. However, since there may not always be a non-communicative outcome signifying the completion of an opinion gap activity, the coders discussed whether such activities are structured to go beyond a mere statement of ideas. For instance, when an activity guides learners to go beyond offering ideas, such as comparing ideas, justifying their reasons and stating whether they agree or disagree, it was accepted to have a communicative outcome.

The discussion session held following the coding process helped the researcher see which types of coding were more sensible and arrive at the final versions of coding before analysing the rest of the activities in the four textbooks for the present study.

Overall, 23 activities were evaluated against eight questions and four criteria for task characteristics, thus making 184 items to be analysed for inter-coder agreement for task characteristics. When all the coders completed marking each activity against eight questions, the analysis data of the researcher and the other coders entered in an Excel spreadsheet was transferred to Statistical Package for Social Sciences (SPSS) to calculate the level of inter-coder agreement (Krippendorff's Alpha). When Krippendorff's Alpha was computed for the analysis of activities in the textbooks based on the eight questions, a high correlation score (.8873) was achieved, which was a satisfactory score to continue coding other activities in the textbooks.

Following the analysis of the intercoder analysis for the task characteristics, the final decisions of the coders for each activity regarding whether they qualify for a task or not needed calculating. Depending on whether an activity meets all four criteria of taskness with positive answers to eight questions for these criteria or not, the activity was coded as a task or non-task (i.e., exercise). In what follows, the relevant data was entered in SPSS for inter-coder agreement. The inter-coder correlations (Krippendorff's Alpha) of the 23 items, representing 23 decisions on taskness were found to be high (.8685). After high correlation rates were achieved for task characteristics and taskness decisions, the researcher continued to code all the other activities in the textbooks accordingly.

Intercoder Reliability Analysis for Q2

After the coders completed the coding of activities as task or non-task, the third inter-coder analysis was conducted with the same coders and decisions were made regarding task types. The researcher trained the coders to analyse the task types. Then, the researcher and the coders examined a few tasks together based on Checklist 2. The types of the tasks to be examined during the training were

purposefully chosen to familiarise the coders with each task type and to have a discussion on the tasks that were open to debate in terms of their type. The training was maintained until the coders reached to point where they felt confident to code tasks separately.

When the coders arrived at an agreement regarding how to make decisions regarding task types, the coders were asked to analyse eight tasks (10 percent of all the tasks in the textbooks examined) against 11 task types separately using the online Checklist 2 on Excel. When the coding session was over, a discussion session was held as in the previous phase of the study to compare and contrast the results and give the coders an opportunity to justify their reasons for the way they coded each task to increase the reliability of coding. A few issues raised during the discussion session are worthy of mentioning.

The first issue was related to the levels of analysis to examine FO tasks (morphological, syntactic, lexical, and pragmatic). Initially, the scope of the focused tasks was mistaken for being limited to grammar rules only by one coder. Then, the researcher provided some additional examples to show that it would also be specific vocabulary or expressions targeted.

The other issues having emerged during the discussion session were about the HN and TT tasks, which are concerned with the presence or absence of the task content. One problem encountered with the coding of tasks with regard to this criterion was that the concept of task input (content) can be mistaken for task instructions. The discussion session revealed that the distinction between task instructions and task content should be underlined. It was decided that while task instructions (e.g., discussion questions) are mainly the directions given to be followed throughout the task implementation process, task input includes the written

or verbal data to be worked on (e.g., a map, a scenario, a role card, an announcement at an airport) for the accomplishment of the task.

The second problem with the decisions made on HN and TT tasks was related to when the task content is present. The main difference between the HN tasks and TT tasks is whether the task-takers can access the task data when implementing the task or not. When a task requires task-takers to search for something and work on the data they have found to accomplish a task, it possesses the characteristics of both task types, for there is no task content provided at the beginning of the process as in TT tasks, but once they find the relevant information online, they can keep it throughout the task as in HN tasks. In these cases, the coders decided to code it as HN task as learners are allowed to revisit task content any time they would like.

Following the discussion session, the analysis data for task types by the researcher and the coders were entered in Statistical Package for Social Sciences (SPSS) to calculate the inter-coder correlations (Krippendorff's Alpha). As a result, a high correlation score (.8178) was achieved for the analysis of the task types. After achieving a high correlation for the analysis, the researcher continued analysing the tasks in the textbooks for their types.

Intercoder Reliability Analysis for Q3

The fourth and final analysis for inter-coder correlations was conducted to check the degree of agreement of the task complexity ratings by the researcher and two independent coders.

As the third question of the study necessitated analysing output-based and integrative tasks, 10 percent of these tasks taken from each textbook were coded by the coders independently after receiving training on task complexity dimensions. That is, seven out of 66 output-based and integrative tasks were analysed against 15

sub-criteria generated for nine task complexity dimensions by the coders based on the task complexity frameworks by Robinson (2001a, 2001b, 2003, 2007, 2008, 2011, 2015) and Skehan and Foster (2001).

Following the coding phase, a discussion session was held to compare and contrast the analysis data from the coders. The discussion session revealed that the coders did not have much problem with the operationalisation of any criterion created for each task difficulty dimension. However, the only problem emanated from deciding whether the task instructions given in the textbooks would make cognitive/conceptual demands resource-directing dimensions are predicted to do. It was especially the case with *causal reasoning* and *intentional reasoning*.

Task-takers are expected to go beyond simple information transformation and “justify beliefs and support interpretations why events follow each other by giving reasons” on the tasks requiring *causal reasoning* (Robinson, 2005, p. 5) and reflect on the “intentions”, “beliefs” and “desires” of individuals on the tasks requiring *intentional reasoning* (Robinson, 2008, p. 10).

The coders stated to have had difficulty in making decisions regarding whether the task instructions in the textbooks would be conducive to the production of certain linguistic features associated with these types of reasoning demands, such as logical subordinators (CR) and psychological and cognitive state terms (IR). After the discussion, the coders agreed that considering that resource-directing dimensions that CR and IR belong to are predicted to make learners push output (Robinson, 2001a), only the tasks with the instructions that make it highly likely for task takers to produce the predicted linguistic features should be coded to require CR or IR since some tasks would simply be accomplished with the mere transfer the information of an event.

After the coding session, Krippendorff's Alpha reliability analysis was computed using the coding data from the researcher and the two independent coders were computed. A high correlation score (.8750) was achieved, which allowed the researcher to continue analysing the rest of the research data.

All in all, the methodology chapter presented the methodology employed in the present thesis. Starting with the discussion of the research design, the chapter continued with the description of the context of the study. Then, it proceeded to the method of sampling and instrumentation. In what follows, the chapter explained how the research data was gathered analysed. Finally, the last two sections of the methodology chapter were devoted to the inter-coder analysis and ethical considerations. The next chapter will be devoted to the results of the current study.

CHAPTER 4: RESULTS

Introduction

The results chapter aims to present the analysis results of the language tasks in four EFL textbooks carried out within the scope of this thesis. This chapter begins with the purpose of this study and the research questions that have been addressed. Next, it provides a brief overview of the research design. Finally, the chapter presents the results for each research question.

This study set out to provide a comprehensive analysis of the tasks in the EFL textbooks used in the Anatolian high school context in Turkey. To this end, it seeks to provide answers to the questions below:

1. To what extent are the activities in the English textbooks used in the Anatolian High school context possess task characteristics?
2. How are task types sequenced in the textbooks?
3. How are task complexity dimensions of output-based and integrative tasks sequenced in the textbooks?

Four English textbooks, designed by the MoNE to be studied in Anatolian high schools in Turkey were chosen to investigate language tasks. Content analysis method was chosen for the analysis of the data in the textbooks. As the first step, all the activities in the textbooks were tabulated in an Excel spreadsheet. To analyse the activities, the checklists to be used for the task characteristics, task types and task complexity dimensions were determined by the researcher based on the theoretical framework. However, coding was performed both inductively and deductively as the analysis process necessitated using inductively created categories or sub-categories

when the data did not fit in the original, primary set of categories deductively created based on the theoretical framework. Afterwards, intercoder analysis was also conducted to ensure the reliability of the coding. In what follows, descriptive statistics (frequencies and percentages) were calculated, and the results were tabulated. Finally, the results were used to compare and contrast the task characteristics, task types and task complexity dimensions found in the textbooks within and across the grade levels.

Results of the Study

In this section, the findings of the study will be presented in a quantitative and descriptive manner. First, the findings related to the task characteristics in the EFL textbooks will be reported. Next, the sequencing of the task types will be revealed. Finally, the sequencing of the task complexity dimensions in the textbooks will be demonstrated.

Tasks and Task Characteristics

As the present study mainly focuses on language tasks as defined by Ellis (2009) and Ellis and Shintani (2014), it was necessary to find out to what extent the activities in the EFL textbooks used in the Anatolian High school context possess the task characteristics prescribed. This step was also important to distinguish tasks from non-tasks (exercises) among the given activities in the textbooks and to exclude the non-tasks from the present study as the study proceeded to the task types and task complexity dimensions in the second and third stages of the analysis. To this end, each activity in the textbooks was evaluated against four criteria of taskness.

Criterion 1

The first criterion utilised in the study was “The primary focus should be on ‘meaning’” (Ellis, 2009, p. 223; Ellis & Shintani, 2014, p. 135). Table 11 demonstrates to what extent the activities in the EFL textbooks satisfy Criterion 1.

Table 11

The Frequencies and Percentages of the Activities Satisfying Criterion 1

Grades	Criterion 1		Total Activity
	<i>f</i>	%	
9	144	44	329
10	136	53	258
11	204	61	334
12	146	63	233
Total	630	55	1154

Table 11 shows that in the textbook for grade 11 has the highest number of activities that satisfy the *meaning* criterion with 204 instances (61%). The other textbooks have pretty similar number of activities meeting this criterion. While 146 (63%) of the activities were discovered to be meaning focused in grade 12, 114 activities (44%) were found in grade 9. Although the meaning-focused activities are relatively less in number (136) in grade 10, they represent 53 per cent of the activities in the book. In total, 630 activities satisfy the criterion, making up the 55 per cent of 1154 activities in four EFL textbooks analysed.

Criterion 2

The second criterion requires “there should be some kind of ‘gap’” (Ellis, 2009, p. 223; Ellis & Shintani, 2014, p. 135). Table 12 shows to what extent the activities in the EFL textbooks meet the criterion.

Table 12

The Frequencies and Percentages of the Activities Satisfying Criterion 2

Grades	Criterion 2		Total Activity
	<i>f</i>	%	
9	128	39	329
10	112	43	258
11	150	45	334
12	94	40	233
Total	485	42	1154

As demonstrated in Table 12, it was found that the textbook for grade 11 has the highest number of activities satisfying the gap criterion, with 150 instances corresponding to 45 per cent of all the activities in the textbook. It is followed by the textbook for grade 9 with 128 activities (39%), the textbook for grade 10 with 112 activities (43%), and, finally, the one for grade 12 with 94 activities (40%).

Criterion 3

According to the third criterion, “learners should largely rely on their own resources (linguistic and non-linguistic) in order to complete the activity” (Ellis, 2009, p. 223; Ellis & Shintani, 2014, p. 135). Table 13 presents the findings regarding Criterion 3.

Table 13

The Frequencies and Percentages of the Activities Satisfying Criterion 3

Grades	Criterion 3		Total Activity
	<i>f</i>	%	
9	190	58	329
10	201	78	258
11	229	69	334
12	190	82	233
Total	810	70	1154

As for the third criterion, it was found that the textbook for grade 11 has the highest number of activities meeting the requirements of criterion 3, with 229 incidences (69%). While 201 activities (78%) meet the criterion in the textbook for

grade 10, 190 activities satisfy this criterion in grades 9 and 12, corresponding to 58 and 82 per cent of the activities in the textbooks, respectively.

Criterion 4

The final criterion, criterion 4, is concerned with whether “there is a clearly defined outcome” (Ellis, 2009, p. 223; Ellis & Shintani, 2014, pp. 135-136). Table 14 demonstrates the findings regarding criterion 4.

Table 14

The Frequencies and Percentages of the Activities Satisfying Criterion 4

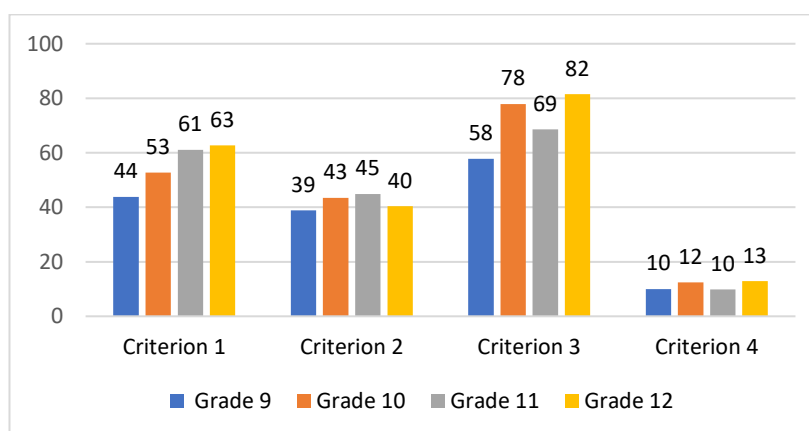
Grades	Criterion 4		Total Activity
	<i>f</i>	%	
9	33	10	329
10	32	12	258
11	33	10	334
12	30	13	233
Total	127	11	1154

When comparing the frequencies and percentages of the activities meeting the fourth criterion within each grade level, it was found that the values in each grade level are quite close to one another (Table 14). In the textbooks for grades 9 and 11, there were found 33 activities, representing the 10 per cent of all the activities in each textbook. Secondly, in the textbook for grade 10, 32 of all the activities (12%) were detected to satisfy the criterion. As for the textbook for grade 12, 30 activities (13%) were identified.

Taken as a whole, the percentages of the activities meeting the task criteria within four EFL textbooks are displayed in Figure 38.

Figure 38

The Percentages of the Activities Satisfying Task Criteria in the EFL Textbooks



It is demonstrated in Figure 38 that more than two-fifths of the activities in each EFL textbook were found to focus on *meaning* (Criterion 1), their percentages increasing as the grade level, and therefore the proficiency level, increases. When it comes to the presence of a communicative *gap* in the activities in the textbooks analysed, it was observed that the percentages in each textbook show similarity. Only around 40 per cent of the activities in each textbook was found to satisfy the *gap* criterion (Criterion 2). As for the percentages of the activities that allow learners to use their own (*non*)*linguistic resources* (Criterion 3), the analysis revealed that more than 50 per cent of all the activities in each textbook comply with this criterion. Finally, as a fraction in the simplest form, only around 1/10 of the activities in each textbook satisfy the last criterion (outcome).

Overall, the results of the present study revealed that the order of the criteria satisfied in each textbook by percentage is the same, criterion 3 (the use of one's own (*non*)*linguistic resources*) being the highest satisfied criterion, followed by criterion 1 (*meaning*), criterion 2 (*gap*) and finally, criterion 4 (*outcome*), being the least satisfied one.

Tasks and Non-tasks

For an activity to qualify for a task, it needs to satisfy all four criteria proposed by Ellis (2009) and Ellis and Shintani (2014). Based on the results of the analysis for characteristics, decisions were made regarding which activities were to be coded as task or non-task (i.e., exercise). Table 15 shows the frequencies and percentages of the tasks and non-tasks identified within four ELT textbooks.

Table 15

The Frequencies and Percentages of the Tasks and Non-tasks in the EFL Textbooks

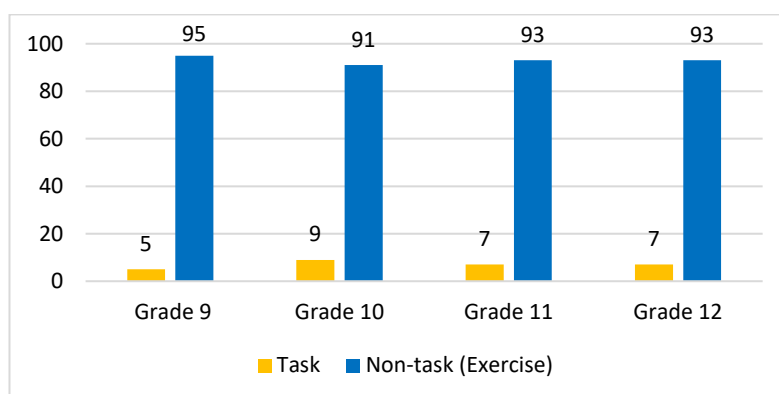
Grades	Tasks		Non-tasks (Exercises)		Total Activity
	<i>f</i>	%	<i>f</i>	%	
9	17	5	312	95	329
10	23	9	235	91	258
11	24	7	310	93	334
12	16	7	217	93	233
Total	80	7	1074	93	1154

In Table 15, the frequencies and the percentages of tasks and non-tasks can be seen for each grade level. In the textbook for grade 9, there were found 329 activities in total with 17 tasks (5%) and 312 non-tasks (95%). In the textbook for grade 10, 258 activities were detected, of which 23 (9%) were tasks and 235 were non-tasks (91%). When it comes to the textbook for grade 11, 24 tasks (7%) and 310 non-tasks (93%) were detected out of 334 activities. Finally, 16 tasks (7%) and 217 non-tasks (93%) were found out of 233 activities in the textbook for grade 12.

It was discovered that the textbook designed for the grade 11 consists of the highest number of tasks (24) of all the textbooks, followed by grade 10 (23), grade 9 (17) and grade 12 (16). As a whole, there were found 80 tasks (7%) and 1074 non-tasks (93%) in a total number of 1154 activities in four EFL textbooks. The percentages of the tasks and non-tasks in the textbooks are displayed in Figure 39.

Figure 39

The Percentages of the Tasks and Non-tasks in the EFL textbooks



From the data presented in Figure 39, it is seen that at least 91 per cent of all the activities in the textbooks consist of non-tasks. On the other hand, only around seven per cent of the activities qualify for a task, as defined by Ellis (2009) and Ellis and Shintani (2014).

When comparing the distribution of tasks and non-tasks (i.e., exercises) within each grade level, it is seen that the values for the percentages of the tasks and non-tasks (i.e., exercises) in each grade level are pretty close to one another. While the textbook for grade 10 has the highest use of tasks with nine per cent, it is followed by grades 11 and 12 with seven per cent of their activities qualifying for tasks, and finally, grade 9 with five per cent.

Task Types and Sequencing of Tasks

The second research question necessitated analysing the task types in four EFL textbooks. Five dual categories based on Ellis's (2003, 2011, 2020) framework were utilised to determine the task types in the textbooks. These categories are (1) *focused* versus *unfocused* tasks, (2) *real-world* versus *pedagogic* tasks, (3) *input-based* versus *output-based* and *integrative* tasks, (4) *here-and-now* versus *there-and-then* tasks and, finally, (5) *open* versus *closed* tasks.

Focused versus Unfocused Tasks

The first category distinguishes tasks based on whether they orient learners to the use of specific linguistic features while not explicitly presenting these features (focused) or allow them to use language communicatively in general (unfocused). The research results regarding the frequencies and percentages of FO or UN tasks in the textbooks for each grade level are given in Table 16.

Table 16

Focused and Unfocused Tasks Identified in the Textbooks

Grades	Focused		Unfocused		Total
	<i>f</i>	%	<i>f</i>	%	
9	17	100	0	0	17
10	21	91	2	9	23
11	20	83	4	17	24
12	14	88	2	13	16

As Table 16 demonstrates, it was found that the majority of the tasks in each textbook were designed to be FO, at least 83 per cent of each targeting the use of certain linguistic features during task implementation. 17 FO tasks (100%) were detected in the textbook for grade 9. When it comes to grade 10, 21 FO tasks (91%) were found. In the textbook for grade 11, 20 FO tasks (83%) were detected. Finally, 14 (88%) of all the tasks in the textbook for grade 12 were found to be FO.

No UN task was found in grade 9. Two were detected in the textbooks for grades 10 and 12, representing 9 and 13 per cent of all the tasks, respectively. Finally, four UN tasks were found in grade level 11, which has the highest per cent of unfocused tasks (17%) among other grades.

Real-world versus Pedagogic Tasks

The second category is concerned with whether a task has situational authenticity and corresponds to real-life events (real-world), or only has interactional

authenticity but still functions to facilitate natural language (pedagogic). The frequencies and percentages of RW and PE tasks are demonstrated in Table 17.

Table 17

Real-world and Pedagogic Tasks Identified in the Textbooks

Grades	Real-world		Pedagogic		Total
	<i>f</i>	%	<i>f</i>	%	
9	2	12	15	88	17
10	5	22	18	78	23
11	4	17	20	83	24
12	4	25	12	75	16

It is presented in Table 17 that the frequencies of RW tasks are pretty close in each textbook, ranging from two to five. The textbook for grade 10 possesses the highest number of RW tasks (five), representing 22 per cent of all the tasks in the textbook. The textbooks for grades 11 and 12 have four RW tasks, corresponding to 17 and 25 per cent of all the tasks in these textbooks, respectively. They are followed by two real-world tasks in grade 9 (12%).

Pedagogic tasks outnumber *real-world* tasks in each textbook. The highest number was found in the textbook for grade 11 with 20 instances of PE tasks (83%). Secondly, 18 PE tasks (78%) were detected in grade 10. It was followed by grade 9 with 15 PE tasks (88%). Finally, grade 12 was found to have the lowest values for frequency and percentage with 12 PE tasks, corresponding to 75 per cent of all the tasks in the textbook.

Input-based versus Output-based and Integrative Tasks

The third category deals with whether a task necessitates only decoding a message and requires no or minimal verbal production (input-based), necessitates oral or written language production (output-based) or entails the use of skills from both input-based and output-based tasks (integrative). The frequencies and percentages of IB, OB and INT tasks are demonstrated in Table 18.

Table 18*Input-based, Output-based and Integrative Tasks Identified in the Textbooks*

Grades	Input-based		Output-based		Integrative		Total
	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%	
9	6	35	3	18	8	47	17
10	4	17	9	39	10	44	23
11	2	8	7	29	15	63	24
12	2	12.5	6	37.5	8	50	16

Table 18 demonstrates that it is possible to find the instances of each task type (i.e., IB, OB and INT) in the four textbooks examined. It is seen that the use of IB tasks shows a decreasing trend in frequency as the grade levels increase. Grade 9 consists of the highest number, with six IB tasks, corresponding to 35 per cent of all the tasks. It is followed by four tasks (17%) in grade 10, two in both grades 11 (8%) and 12 (12.5%).

When it comes to OB tasks, the highest number of this type of tasks was identified in the textbook for grade 10 with nine incidences, representing 39 per cent of all the tasks in the textbooks. Next, seven tasks were detected in the textbook for grade 11 (29%). It was followed by six OB tasks (37.5%) in grade level 12 and three (18%) in grade level 9.

As for the third category, INT tasks, the analysis revealed that their frequencies range from eight to 15, with eight INT tasks in grades 9 (47%) and 12 (59%), 10 tasks in grade 10 (44%), 15 tasks in grade 11 (67%).

Here-and-now versus There-and-then Tasks

When it comes to the fourth category, it investigates if learners are allowed to see the task-related content in front of them (here-and-now) or they have no access to it (there-and-then) during task implementation. The results of the study regarding the HN and TT tasks are provided in Table 19.

Table 19*Here-and-now and There-and-then Tasks Identified in the Textbooks*

Grades	Here-and-now		There-and-then		Total
	<i>f</i>	%	<i>f</i>	%	
9	11	65	6	35	17
10	20	87	3	13	23
11	18	75	6	25	24
12	9	56	7	44	16

It is shown in Table 19 that the textbook for grade 10 has the highest number of HN tasks with 20 instances (87%). It is closely followed by grade 11 with 18 HN tasks (75%). While grade 9 has 11 (65%) HN tasks, grade 12 has 9 (56%).

As for the TT tasks, their frequencies range from three to seven in the textbooks analysed. While seven TT tasks (44%) were detected in the textbook for grade 12, six were identified in grades 9 and 11, corresponding to 35 and 25 per cent of all the tasks in these textbooks, respectively. Finally, only three TT tasks (13%) were observed in the textbook for grade 10.

Closed versus Open Tasks

The last dual category deals with the number of task outcomes. While some tasks can have a single or limited number of outcomes (closed), some may have several different outcomes (open). Table 20 demonstrates the descriptive results regarding task outcomes (CL versus OP).

Table 20*Closed and Open Tasks Identified in the Textbooks*

Grades	Closed		Open		Total
	<i>f</i>	%	<i>f</i>	%	
9	8	47	9	53	17
10	8	35	15	65	25
11	7	29	17	71	24
12	7	44	9	56	16

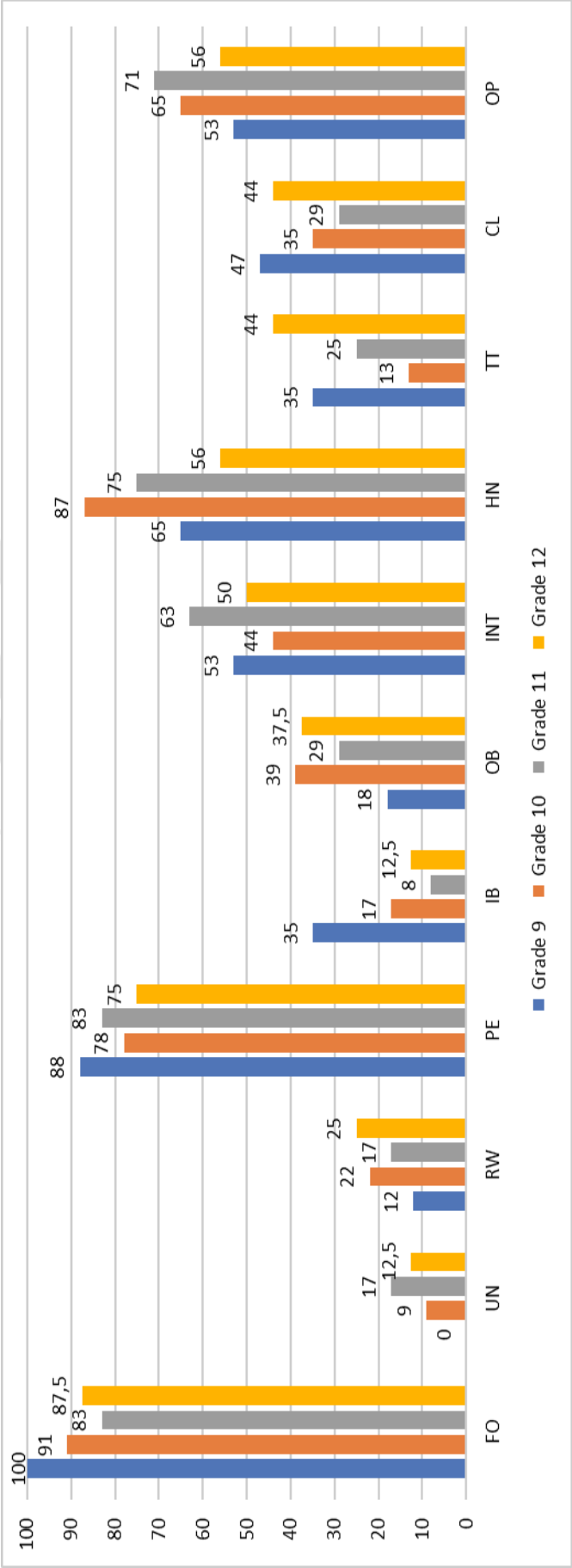
The analysis revealed that the frequencies of CL tasks are pretty similar in each textbook (Table 20). Grades 9 and 10 have eight instances of CL tasks, corresponding to 47 and 35 per cent of all the tasks in the textbooks, respectively. The textbooks for grades 12 and 11 consist of seven instances of CL tasks, representing 44 and 29 per cent of all the tasks in these textbooks, respectively.

As for the OP tasks, 17 instances of them were identified in grade 11 (71%). Secondly, 15 OP tasks (65%) were found in grade 10. When it comes to the textbooks for grade levels 12 and 9, the results are pretty close in frequency and percentage. Both have nine OP tasks, and the percentages of these tasks in the textbooks are 53 (grade 9) and 56 (grade 12).

Taken as a whole, the percentages of the task types identified within four EFL textbooks are displayed in Figure 40.

Figure 40

The Percentages of the Task Types in the EFL Textbooks



Note. FO = focused, UN = unfocused, RW = real-world, PE = pedagogic, IB = input-based, OB = output-based, INT = integrative, HN = here-and-now, TT = there-and-then, CL = closed, OP = open.

It is clearly seen in Figure 40 that at least 80 per cent of the tasks in the textbooks examined consist of FO tasks, while only a small proportion of the tasks are UN. Similarly, more than 50 per cent of the activities qualify for PE, HN and OP tasks, being commonly used in the textbooks as opposed to their counterparts, RW, TT and CL tasks. When it comes to the categorisation based on IB versus OB and INT tasks, it is seen that OB and INT tasks together outnumber IB tasks in each textbook. There was not detected an increasing or decreasing trend across the grade levels in terms of any task types given.

Task Complexity and Sequencing of Tasks

The third research question of this study focuses on the task complexity dimensions of output-based and integrative tasks and how they are sequenced within units. Nine task complexity dimensions based on Robinson's (2001a, 2001b, 2003, 2007, 2008, 2011, 2015) and Skehan and Foster's (2001) frameworks were examined in the study. The task complexity dimensions include two main categories: resource-directing and resource-dispersing dimensions. The sub-categories of resource-directing dimensions examined in the present study are (1) *there-and-then* (TT), (2) *spatial reasoning* (SR), (3) *causal reasoning* (CR), (4) *intentional reasoning* (IR) and (5) *perspective-taking* (PT). Secondly, the sub-categories of resource-dispersing dimensions analysed in this study are (1) *single task* (ST), (2) *task structure* (TS), (3) *single step* (SS) and (4) *prior knowledge*.

According to Robinson (2015), cognitive task demands (i.e., task complexity dimensions) should be taken as a basis for task sequencing and tasks should be made more complex step by step, starting tasks simple in both resource-directing and resource-dispersing dimensions (*stable (SS) attractor* stage), moving on to the tasks made complex in resource-dispersing dimensions (*automatisation (A)* stage) and

finally ending with the ones complex in both dimensions (*restructuring (R)* and *complexity (C)* stage). This section is devoted to the findings related to the analysis of OB and INT tasks based on the task complexity dimensions and the principles of SSARC.

Resource-directing Dimensions

Resource directing dimensions are related to the conceptual/communicative demands tasks pose (Robinson, 2015). The findings related to the resource-dispersing dimensions (there-and-then, spatial reasoning, causal reasoning, intentional reasoning and perspective-taking) are demonstrated in Table 21 below.

Table 21

Resource-directing Dimensions Identified in the Output-based and Integrative Tasks

Grade	Resource-directing Dimensions									
	(+) TT		(+) SR		(+) CR		(+) IR		(+) PT	
	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%
9	4	36	1	9	0	0	2	18	1	9
10	5	26	1	5	3	16	0	0	2	11
11	15	68	0	0	4	18	3	14	5	23
12	3	21	0	0	4	29	1	7	2	14

Note. TT = there-and-then, SR = spatial reasoning, CR = causal reasoning, IR = intentional reasoning, PT = perspective-taking.

It is seen in Table 21 that the number of tasks manipulated along the *there-and-then* (TT) dimension ranges from three to 15. The textbook for grade 11 consists of the highest number of TT tasks with 15 instances (68%). It is followed by five TT tasks (26%) in grade 10, four (36%) in grade 9 and three in grade 12 (21%).

As for the *spatial reasoning* (SR) dimension, only one task in the textbooks for grades 9 and 10 is made complex along this dimension, corresponding to 9 and 5 per cent of the OB and INT tasks in those textbooks, respectively. Nevertheless, no tasks were manipulated along this dimension in grades 11 and 12.

Third, while no tasks were made complex in *causal reasoning* (CR) dimension in the textbook for grade 9, three tasks (16%) were manipulated along this dimension in grade 10 and four tasks in grades 11 (18%) and 12 (29%).

When it comes to the *intentional reasoning* (IR) dimension, while no tasks were found manipulated along this dimension in the textbook for grade 10, one task (7%) was detected in grade 12, two (18%) were detected in grade 9 and three (14%) in grade 11.

Finally, the number of tasks necessitating *perspective-taking* (PT) ranges from one to five in the textbooks examined. The highest number of tasks manipulated along this dimension belongs to grade 11 with five incidences (23%), followed by grades 12 (14%) and 10 (11%) with two instances, and finally, grade 9 (9%) with one task.

Resource-dispersing Dimensions

Resource-dispersing dimensions are related to the performative and procedural demands tasks make on cognition (Robinson, 2007, 2010, 2011, 2015). The findings related to the resource-dispersing dimensions (single task, task structure, single step and prior knowledge) are presented in Table 22.

Table 22

Resource-dispersing Dimensions Identified in the Output-based and Integrative Tasks

Grade	Resource-dispersing Dimensions							
	(-) ST		(-) TS		(-) SS		(-) PK	
	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%
9	8	73	1	9	9	82	0	0
10	6	32	0	0	13	68	2	11
11	13	59	2	9	18	82	3	14
12	10	71	1	7	5	36	1	7

Note. ST = single task, TS = task structure, SS = single step, PK = prior knowledge.

It is demonstrated in Table 22 that the textbook for grade 11 has the highest number of tasks that requires performing more than one action concurrently (-ST) with 13 instances (59%). It is followed by the textbook for grade 12 with 10 tasks (71%), grade 9 with eight tasks (73%) and finally, grade 10 with six tasks (32%).

Secondly, only a few instances of tasks that do not provide task-takers with clear 'task structure' for successful task performance were found in the textbooks. While no tasks were identified to lack clear task structure in the textbook for grade 10, only one was found in the ones for grade 9 and 12, corresponding to the 9 and 7 per cent of the OB and INT tasks in the textbooks, respectively. Finally, two tasks (9%) in the textbook for grade 11 were found not to present a sufficient task structure.

As for the number of task steps, the analysis revealed that the textbook for grade 11 has the highest number of tasks with more than one step (-SS) with 18 instances (82%). Secondly, there were found 13 tasks (68%) with -SS condition in the textbook for grade 10. When it comes to the textbook for grade 9, nine tasks with -SS condition (82%) were identified. Finally, five tasks (36%) were detected in grade 12.

Prior knowledge was the last condition in the resource-dispersing dimensions explored. It was detected that no tasks were presented without sufficient data to make sure learners have prior knowledge to accomplish the task goals in the textbook for grade 9. Merely one task (7%) fails to provide sufficient prior knowledge (-PK) in grade 12, two (11%) in grade 10 and three (14%) in grade 11.

SSARC Steps

Once the tasks in the textbooks were examined based on the resource-directing and resource-dispersing dimensions, the next step was to determine whether

the OB and INT tasks in the same units were sequenced to increase complexity in line with the principles of SSARC by Robinson (2015). The results related to task complexity dimensions according to SSARC steps are displayed in Table 23.

Table 23

Task Complexity Steps Identified in the Output-based and Integrative Tasks Based on the SSARC Model

Grade	(-) RDir (-) RDis		(+) RDis		(+) RDir		(+) RDir (+) RDis	
	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%
9	0	0	5	45,5	1	9	5	45,5
10	4	21	8	42	0	0	7	37
11	2	9	5	23	0	0	15	68
12	2	14	6	43	1	7	5	36

Note. RDir = resource-directing, RDis = resource-dispersing.

The frequencies and the percentages of tasks increased in task complexity dimensions in each grade level can be seen in Table 23. The first and the simplest step (-RDir -RDis) stands for the *stable attractor* (SS) stage in the SSARC model and indicates that a task is not manipulated along either of the dimensions. It is seen that there are not any tasks in the SS stage in the textbook for grade 9. On the other hand, two instances of the tasks in this stage were found in grades 11 and 12, representing 9 and 14 per cent of all the OB and INT tasks in these grades, respectively. Finally, four tasks in SS stage (21%) were found in grade 10.

The second step in the SSARC model is to increase task complexity along the resource-dispersing dimensions (+RDis), and this is called *automatisation* (A) stage. It was detected that five of the tasks in grades 9 and 11 were manipulated only along the resource-dispersing dimension(s), representing 45,5 and 23 per cent of all the OB and INT tasks in these grades, respectively. While six tasks (43%) were found to be in this step in grade 12, eight (42%) were found in grade 10.

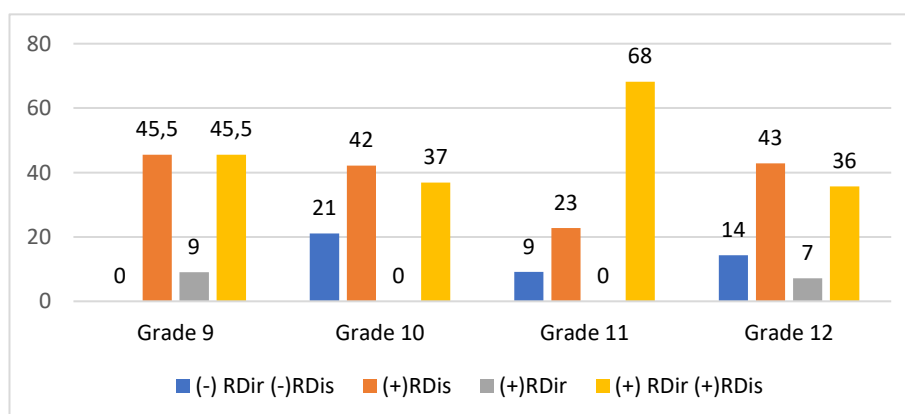
Although there has not been specified a stage in the SSARC model where only resource-directing dimensions are made more complex while other dimensions are kept simple, Table 23 demonstrates the frequencies and percentages of the tasks manipulated only along these dimensions. One task in the textbooks for grades 9 and 12 was found to be complex only in resource-directing dimensions, corresponding to 9 and 7 per cent of the OB and INT tasks in these textbooks, respectively.

In the third step of SSARC, tasks are made complex in both dimensions (+RDir +RDis), and it is named *restructuring and complexity* (RC) stage. It is clear from Table 23 that the highest number of tasks manipulated along both dimensions can be found in the textbook for grade 11 with 15 instances (68%). It is followed by grade 10 with seven tasks (37%) and grades 9 and 12 with five tasks, representing 45.5 and 36 per cent of the tasks in the textbooks for these grades, respectively.

In the present study, it was aimed to find out whether the tasks in the units were manipulated to be more complex in a principled way in line with the principles of SSARC. Taken as a whole, the percentages of the task complexity steps based on the SSARC model within four EFL textbooks are displayed in Figure 41.

Figure 41

The Distribution of Task Complexity Steps Based on the SSARC Model



Note. RDir = resource-directing, RDis = resource-dispersing.

It is demonstrated in Figure 41 that the percentages of the OB and INT tasks in different steps of the SSARC model show differences across grade levels. It was noticed that, in each grade level, most of the OB and INT tasks are at either step 2 (A) represented as +RDis, or step 3 (RC) represented as +RDir +RDis in Figure 41. The tasks at the simplest step (SS), represented as +RDir +RDis in Figure 41, were the third most used tasks in grades 10, 11 and 12, while they were the tasks that were made complex along only -RDir stage for grade 9. The least common stage was the simplest step (SS) for grade 9, while it was the step that posed only resource-directing demands (+RDir) for the other grade levels.

It must be noted that since the SSARC model has been developed to make recommendations regarding the sequencing of tasks from simple to complex, the distribution of these stages within the units of each textbook was examined for each type of task. Nevertheless, no systematic progression from simple to complex tasks as in the SSARC model was detected in the EFL textbooks examined in the present study.

All in all, the results chapter presented the results of the analyses conducted for the questions of the present study. Starting with a brief discussion of the research purposes, questions and design, the chapter continued with the findings related to the task characteristics in the ELT textbooks, sequencing of the task types as well as the sequencing of the cognitive complexity dimensions of the OB and INT tasks in the EFL textbooks examined. The next chapter will be devoted to the discussion of the research findings.

CHAPTER 5: DISCUSSION

Introduction

The chapter aims to discuss the findings of the present study. It starts with the overview of the research design that summarises the purpose of the present study, the instrumentation and the analysis conducted. Next, it moves on to the discussion of the main findings of the current research concerning the task characteristics, task types and task complexity dimensions in terms of their contribution to the related literature. Then, it continues with the implications for practice and further research. Finally, it ends with the limitations and delimitations of the present study.

Overview of the Study

The present thesis was an attempt to explore the language tasks within the EFL textbooks used in the Anatolian high school context in Turkey. To this end, the study aimed to address the following research questions:

1. To what extent are the activities in the English textbooks used in the Anatolian High school context possess task characteristics?
2. How are task types sequenced in the textbooks?
3. How are task complexity dimensions of output-based and integrative tasks sequenced in the textbooks?

Four EFL textbooks designed by a local community in MoNE and used for grades 9, 10, 11 and 12 during the education year of 2020-2021 were investigated through purposeful selection to gain insights into the language tasks used in the Anatolian high schools in Turkey. The data were accessed from the Educational Informatics Network (EBA) website provided by MoNE and investigated through content analysis.

To answer the questions of the study, the data in the textbooks was analysed by using the checklists prepared with a primary set of categories based on relevant frameworks for task characteristics (Ellis, 2011; Ellis & Shintani, 2014), task types (Ellis, 2003, 2011) and task complexity dimensions (Robinson (2001a, 2001b, 2003, 2007, 2008, 2011, 2015; Skehan & Foster, 2001) in Excel spreadsheet. Intercoder analysis (Krippendorff's Alpha) was conducted for each step of coding to ensure the reliability of the analysis. Once the data analysis was finalised, descriptive statistics (i.e., frequencies and percentages) were conducted, and the results were tabulated.

Discussion of Major Findings

Tasks and Task Characteristics

The first question of the study aimed to find out to what extent the activities in the EFL textbooks used in the Anatolian High school context possess the task characteristics as determined by Ellis (2009) and Ellis and Shintani (2014). To that end, the first step of the data analysis was devoted to identifying the task characteristics the activities in the textbooks possess.

Before discussing the task characteristics identified in the activities in the textbooks examined, it is worthy of note that there is a significant difference in the number of activities in these textbooks. While there are around 330 activities in the textbooks designed for grades 9 and 11, there can be found around 240 in grades 10 and 12, which considerably affects the study results. The difference can be rooted in the fact that the textbooks for grades 9 and 11 and grades 10 and 12 were written by almost the same groups of writers. Therefore, it should also be considered when interpreting the study findings.

The relevant results demonstrated that most of the activities in the textbooks examined meet the first and third criteria (i.e., (1) meaning and (3) use of one's own

(non) linguistic resources) for identifying tasks, the majority of them were not counted as ‘task’ since they generally fail to meet the second and the fourth criterion, (i.e., (2) gap and (4) outcome) (Ellis, 2011; Ellis & Shintani, 2014). Specifically, it was found that at least half of the activities in each textbook satisfy criterion 1 (i.e., the focus on ‘meaning’) and the proportion of the activities in the textbooks meeting criterion 3 (use of one’s own (non)linguistic resources) is relatively much higher. On the other hand, while less than half of the activities satisfy criterion 2 (gap), only around one-tenth of the activities meet criterion 4 (outcome) as a fraction in simplest form. The implications of the results regarding each criterion will be further discussed in this section.

As for criterion 1 (meaning), the results show a trend toward an increase in the proportion of meaning-focused activities as the grade levels increase. This finding indicates that learners are more supported to be engaged in communicative processes and oriented toward using L2 as a tool, rather than an end itself as their proficiency levels increase.

Concerning criterion 1, it is important to note that some of the activities were marked as non-task as they fail to engage learners in decoding/encoding pragmatic meaning (“the highly contextualised meanings that arise in acts of communication”), although they orient learners to semantic meaning (“the meanings of lexical items or of specific grammatical structures”) (Ellis, 2005, p. 12). Many of the activities coded as non-task are essentially meaning-focused; however, they only deal with linguistic/surface meaning of the words, neglecting the functional/intentional meaning they would communicate in context.

Considering the results regarding criterion 2 (gap), they demonstrate that more than half of the activities in the textbooks examined lack a communicative gap

or any sort of gap to be closed, which would motivate learners to use L2 to communicate to close it. Some of the most commonly encountered activities failing to satisfy this criterion include matching activities (e.g., word-meaning, sentence-picture), comprehension questions about a recording/text, writing a paragraph or preparing a speech without an audience to address as well as rewriting or retelling a story when all the other learners already know the story and so on.

Interestingly, although the textbooks are comprised of several real-life activities with communicative outcomes (e.g., e-mail, text message, poster, magazine article, comment on a blog, letter of complaint), they were considered non-task since there was no transfer of information from one person to another, one form to another or one place to another. Since these tasks demonstrate some of the task characteristics, they can be considered to be *task-like* as they are named in previous studies (Jaruteerapan, 2020; Terekhova, 2017)

The lack of communicative *gap* in the activities would negatively affect the use of L2. For instance, some of the activities mentioned above fail to create context, in which there is a need for communication in L2 as learners will be aware that there is no need to transfer information. In this case, the activity does not go beyond practising L2 for its own sake as in exercises rather than as a means for communication.

Another reason why some activities are considered to fail to introduce a gap is that it is not clear from the task instructions in which way the transfer of a message was expected to occur. For instance, some activities have only questions without task instructions, which does not inform whether these questions will be discussed or written to be shared afterwards. It is also ambiguous if the transfer of information (if ever takes place) would be from students to students or students to teacher.

One of the most important findings to emerge from the research data is that the use of one's own *(non)linguistic resources* (criterion 3) is the most satisfied criterion in each textbook. This result underlines that priority is placed on giving learners a choice over the resources to be utilise rather than teaching or dictating specific linguistic features to be focused on. This can be taken as evidence that even if non-tasks outnumber tasks in the textbooks, the activities in the textbooks mainly encourage production and creative use of L2 rather than controlled and guided practice as in non-tasks (i.e., exercises).

The most striking result to emerge from the data is that almost all of the activities in each textbook fail to present a communicative (e.g., non-linguistic) *outcome* that goes beyond the mere demonstration of linguistic knowledge and signals the end of a task implementation process. This makes it difficult to distinguish the linguistic performance from the task outcome, which requires using L2 as a tool. This result, therefore, offers compelling evidence for L2 practice's being taken as an end itself in the majority of the activities analysed. That many of the activities meeting the first three task criteria were marked as non-task since they fail to have a 'communicative outcome' can be interpreted that criterion 4 stands as the most decisive factor of all to distinguish tasks from non-tasks in these textbooks.

As for the outcomes of the tasks detected in the textbooks in the current study, they can be grouped under three types of communicative outcomes as classified by Jaruteerapan (2020). They include "observable results" (p. 72) such as a completed chart/diagram, route drawn on a map, written slogans/mottos, lists/notes, ordered texts/dialogues/pictures, PowerPoint presentations, blog entries, wish tree, video, paragraphs/essays/application, letters/reports/stories/magazine articles, and travel guides/booklets. The second group of outcomes, an "abstract idea" (p. 72),

include finding the source of problems, proposing solutions for problems, decision-making, debating, ranking, and role plays for such actions as enrolling on a fitness centre and making a reservation. Finally, the third group of outcomes include “physical reactions” (p. 72) such as labelling something or someone by listening to a recording.

Considering that less than half of the activities satisfy criteria 2 and 4 in the textbooks examined in this study, the results of the study are in line with the previous findings that *gap* and *outcome* are often the missing criteria in the activities designed (Jaruteerapan, 2020; Peng & Pyper, 2021). Our results also share a similarity with the findings in Erlam’s (2016) study, where it was found that the tasks designed by teachers mainly failed to satisfy criterion 2 (*gap*) and criterion 3 (the use of one’s own (non)linguistic resources). Moreover, the results of the present study regarding the *outcome* criterion are in line with the findings of the study by Baklacı (2018). The study (Baklacı, 2018) revealed that the language activities in the language textbook mostly failed to satisfy the criterion, which is having a meaningful purpose rather than the display of the language items learned before.

The investigation of task characteristics in the EFL textbooks shows that these textbooks mainly consist of non-tasks, while tasks are much fewer in percentage on average. When the textbooks of each grade level are compared, the proportions of tasks and non-tasks (i.e., exercises) in each textbook are considerably similar to one another, indicating that each textbook follows a similar format in terms of the allocation of tasks and non-tasks. As only a small number of the activities were found to qualify for a task, the results of the present study also confirm the previous research results, which demonstrated that non-tasks (i.e., exercises)

outnumber tasks in ELT textbooks (Barrot, 2019; Butler et al., 2018; Dao & Newton, 2021; Günay, 2007; Terekhova, 2017).

The dominance of exercises vis-à-vis tasks in the textbooks analysed within the scope of the present study can be interpreted in line with Ellis's conclusion that popular language coursebooks still consider language as an object with its particular linguistic forms and their meanings contextualised through listening and reading activities (2016). This finding is in good agreement with the previous analysis of the activities in ELT textbooks, which revealed that the textbooks mainly consist of workplans focusing on text manipulation (i.e., exercises) rather than the ones that focus on text-creation (i.e., tasks) (Ellis, 2016; Terekhova, 2017).

It should be noted that even though non-tasks (i.e., exercises) outnumber tasks in the textbooks investigated in the present study, a considerable majority of the activities were coded as non-task since they fail one or two criteria of taskness based on Ellis's framework (Ellis, 2003; Ellis & Shintani, 2014; Jaruteerapan, 2020). The activities in the textbooks examined within the scope of this study do not, however, genuinely consist of traditional, mechanical exercises that mainly focus only on form drawing on the behaviourist approach to language learning since they still satisfy some of the task criteria (e.g., (1) meaning and (2) the use of one's own (non)linguistic resources).

It is also important to note that there was not a significant difference between the textbook designed for grade 9 (A1/A2) and the others for grades 10, 11 and 12 in terms of the percentages of tasks and exercises. In this regard, the findings of the present study differ from the observation of Shintani (2012, p. 217) that "intentional language learning" is adopted in the language textbooks targeting beginner-level

learners since almost half of the activities in the textbooks for grade 9 in the present study were found to be meaning-focused and present a communicative gap.

In the textbooks, there are several examples of task-like activities that bear some of the task features but were considered non-task due to several factors such as the wording in the instructions of the activities, the way activity samples are provided as well as the absence of specific non-linguistic outcomes to emerge as a result of task performance. To illustrate, each textbook consists of role-play activities that would constitute good examples of tasks if their instructions had been written differently. Most of these activities, however, are presented with instructions that explicitly orient learners to specific forms, ask them to recycle some pre-taught linguistic elements or fail to provide them with a well-defined outcome to be achieved. Due to these factors, these activities orient learners to study specific linguistic forms and their functions rather than using L2 as a tool for communication. The findings of the study, therefore, support the argument that even small variations in the activity workplans can change the way languages are learned (Ellis, 2003).

Although it is claimed in the 9th-12th *Grades English Curriculum* that the activities within the textbooks examined in the present study were designed to be “task-based, collaborative, and project based” (MoNE, 2018, p. 6), the percentages of the tasks and non-tasks in the textbooks indicate that these textbooks do not essentially adopt a *task-based syllabus* as tasks do not constitute the primary units of content in the syllabi (Ellis et al., 2019; Révész, n.d.).

The syllabi can be considered as instances of task-supported syllabus also because tasks in the textbooks still play an important role there even if they are not the primary syllabus component in line with what Révész (n.d.) poses. Besides, the existence of the other components that define those syllabi and assessment, such as

themes, functions and vocabulary, shows that the textbooks investigated adopt a task-supported approach (Révész, n.d.). In addition, the syllabi of these textbooks do not comprise unfocused tasks (i.e., the tasks that enable learners to use language in general communicatively), which are primarily used as a part of task-based language teaching (Ellis, 2009). Overall, it is important to note that the textbooks examined within the scope of this study do not actually pursue a task-based approach and progression in their overall design.

This finding supports the argument that although publishers often claim the language textbooks that they design are *task-based* (Littlewood, 2004), they are, in fact, not (Günay, 2007; Long, 2016) or, at best *task-supported* (Butler et al., 2018; Ellis, 2018; Newton & Bui, 2017; Révész, n.d.; Samuda & Bygate, 2008). This finding confirms the previous findings that textbook writers differ in their understanding of tasks and design (Barrot, 2019; Günay, 2007). It also confirms the findings of the previous queries into language textbooks in that the difference between *task-based* and *task-supported* learning has not been clearly understood in the field of L2 language teaching (Ganta, 2015).

Last but not least, it is important to note that the focus on tasks in the present study does not suggest that non-tasks (i.e., exercises) are of no value and should be excluded entirely from the language textbooks. They are also valuable means to study a language (Ellis, 2013, 2018). Besides, non-tasks can be practical in several ways, such as preparing learners for the tasks to be performed subsequently.

Task Types and Sequencing of Tasks

The second step of the data analysis dealt with the five dual-categories of tasks based on the task typology framework by Ellis (2003, 2011, 2020). These categories consist of (1) *focused* versus *unfocused* tasks, (2) *real-world* versus

pedagogic tasks, (3) *input-based* versus *output-based* or *integrative* tasks, (4) *here-and-now* versus *there-and-then* tasks and, finally, (5) *open* versus *closed* tasks.

The first dual category consists of *focused* and *unfocused* tasks. The textbooks in the present study seem to adopt syllabus composed of both FO and UN tasks, even if they are not evenly distributed. When the frequencies and percentages of FO and UN tasks were analysed, it was found that most of the tasks in each grade are FO tasks. It was also revealed that while the percentages of FO tasks show a decline from grade 9 to 11, the percentages in grade 12 are higher than in grade 11.

Interestingly, it was found all the tasks in grade 9 are FO tasks, targeting A1+A2 level learners, although these tasks are not suggested to be used with the low proficiency level learners (Ellis, 2020). One reason the textbook for grade 9 consists of a higher proportion of FO tasks could be that there are a higher number of IB tasks in this textbook than in the other textbooks. It is relatively easier for task designers to incorporate the target linguistic features into the task content and to direct learners' attention to target features through the IB tasks as learners have no option but to focus on these features to accomplish the task outcomes.

Focused tasks have three types: "structure-based production task", "interpretation task" and "consciousness-raising task" (Ellis, 2003, p. 151). It was found that almost all of the FO tasks in the coursebooks are *structure-based production tasks*, whereas only two qualify for *consciousness-raising task* and none for *interpretation task*.

As a matter of fact, it was discovered that the textbooks designed for 10th and 12th graders include *Let's Explore* sections, where students are expected to engage in activities that focus on form. These sections also bear some features of interpretation and consciousness-raising tasks since the activities in these sections are mainly

discovery-oriented (i.e., they do not explicitly teach grammar) and have consciousness-raising functions. This finding supports the previous findings that discovery-oriented activities have become commonly used in English language teaching textbooks recently (Ellis, 2016). Nevertheless, these activities were not counted as a FO task in this study as they consist of solely discrete-point tests with multiple choice questions, which are non-communicative (Shintani, 2016). Besides, the activities coded as *consciousness raising task* in the current study were actually in different sections of the units within the textbooks.

When it comes to the findings regarding *real-world* and *pedagogic* tasks, it was suggested in the literature that PE tasks are employed at the early stages of language learning until learners become capable of deal with full version of a target task (Long, 1985). For the same reason, PE tasks are recommended to precede RW tasks and are practised with lower proficiency levels (Ellis, 2020). Nevertheless, both types of tasks are presented in each textbook examined in the current study.

Although grade 12 has the highest score for RW tasks by percentage as suggested, there is no gradual increase in neither the frequency nor the percentage of RW tasks as grade levels, and therefore the proficiency levels, increase.

One issue related to RW tasks that emerged during the analysis process is worthy of consideration. Although some tasks in the textbooks analysed were coded as RW tasks in the present study, it is questionable whether the tasks can truly qualify for RW tasks. For instance, learners are frequently asked to write a letter (e.g., a letter of complaint) in the textbooks. Although writing a letter is a task people used to perform in natural contexts outside the school in the past, it may not correspond to the tasks the target age group of these textbooks (i.e., aged between 14-18) would generally feel the need to deal with on a daily basis. Even if the genre

of letter has not yet become completely obsolete, it may not be more than a PE task if it continues to be presented in the textbooks with the generations to come as it has long started to lose its situational authenticity.

As for the input-based, output-based and integrative tasks in the textbooks, there was no increasing or decreasing trend with the grade level change. The analysis of the IB tasks in the textbooks revealed that the sequencing of these task types shows partial agreement with the relevant suggestions made in the literature. The findings regarding the IB tasks differ from the suggestions by Prabhu (1987), Ellis (2020) as well as Ellis et al. (2019) to some extent. It is expected task based instruction starts with IB tasks at lower proficiency levels and moves on to OB tasks as proficiency levels increase. In the present study, however, it was found that the majority of the tasks in grade 9 are not IB. Moreover, although IB tasks show a decrease by percentage as grade levels increase from grade level 9 to 11, their percentage increases in grade level 12 again. The findings, nevertheless, lend support to the observation by Shintani (2012, p. 271) that the textbook targeting beginner-level learners (grade 9) adopt a “production-based approach” to language learning as the majority of the tasks in this textbook are OB or INT tasks.

As for the dual categories of *here-and-now* and *there-and-then* tasks and *closed* and *open* tasks, no consistent results were found that would indicate an increasing or decreasing trend among the grade levels. It is worthy of note that HN tasks outnumber TT tasks in each textbook, meaning that learners are generally provided with or allowed to check the task content whenever they would like during the implementation of the tasks in the textbooks. Likewise, OP tasks exceed CL tasks in number in all the textbooks. It may be interpreted that the majority of the tasks in the textbooks were designed to have several possible outcomes, namely solutions.

Considering the percentages of the task types across grade levels, it can be claimed that the sequencing of the task types among the textbooks examined is not in line with the recommendations by Ellis (2020) for task type choice for different proficiency levels. Ellis (2020) suggests that certain types of tasks are more suitable for particular groups of learners at certain language levels, and there should be an increasing focus from *pedagogic, input-based, here-and-now*, and *closed* tasks to *real-world, focused, output-based, there-and-then* and *open* tasks with the increase in the grade levels.

In general, there was not a smooth, consistent transition from lower grade levels to the higher ones regarding the task types employed. One reason could be that task type choices in these textbooks might have been affected by the preferences of the writer groups designing the textbooks. It was aforementioned that the textbooks for grades 9 and 11 and grades 10 and 12 were written by almost the same groups of writers. As the task types to be included in the textbooks can be affected by factors such as the beliefs and approaches of the textbook writers (Aleml, 2014), this fact should also be considered when interpreting the findings for the sequencing of the task types across grade levels.

It is worth mentioning that the task type analysis process revealed that even though the dual classification of task types is practical for making decisions when designing and sequencing tasks for EFL textbooks, the distinction between the task types in each category is not clear-cut. A few issues arose when analysing the tasks in the textbooks that should be considered when analysing task types.

It was challenging to make decisions regarding the authenticity of the tasks in the textbooks. There were instances of some tasks whose instructions bear the characteristics of both pedagogic and real-world situations, thus making it difficult to

distinguish one from another. For example, the task instructions can ask a learner to read a text to find out whether their daily life habits are healthy or not. Although it may initially seem to be a pedagogic task as learners are generally provided with questions in the pre-reading stage so that they get motivated to read a text in the school context, the same group of learners may read a very similar article with a similar motivation from digital sources such as Buzzfeed or Reddit for non-pedagogical purposes. In this regard, this observation made in the present study confirms Nunan's (1989) argument that real-world or pedagogic tasks should be taken as a continuum rather than two distinct types.

As for the dual category of focused versus unfocused tasks, the decision-making process was not straightforward. Even if the contents, objectives and outcomes of each unit in the textbooks can be visited to see what linguistic features are targeted in a textbook, it is still challenging to find out whether these features would be *very useful*, *must* or highly likely to *naturally occur* during task implementation as predicted by Loschky and Bley-Vroman (1993), especially when OB tasks are in question unless they are interpretation or conscious-raising tasks, which are relatively easier to detect since they directly subject learners to the target feature.

The second issue that emerged when working on focused tasks concerns the levels of analysis. In the present study, the researcher set off on the data analysis by examining whether certain linguistic features were intended to be practised at morphological, syntactic, lexical, and pragmatic levels. However, it was observed that certain language functions were studied in the textbooks by being incorporated into specific genres. When learners were implicitly oriented to practise the predetermined functions (pragmatics level) to perform some tasks (e.g., complaining,

apologising inviting), they were also expected to comply with some conventions of the genres in question (e.g., a CV) to properly accomplish the task goals including but not limited to style, greetings and signoffs, proper grammar and vocabulary and even some organisational specifications. In the present study, the morphological, syntactic, lexical, and pragmatic levels were taken into account; however, the textbook analysis also raised the question of whether the concept of FO task can also include genre-specific features targeted.

Since organisational specifications of different genres are closely related to linguistic features used, it makes one wonder whether multimodality should also be considered in focused tasks. Studying certain genres encompasses some other elements in addition to linguistic features, which go hand in hand with these features for successful task performance. Take a task which asks learners to prepare a traveller's guide about a city they have visited. For this task, learners do not only focus on how to inform people but also pay attention to their tone, style, linguistic choices, text organisation or even the choice of visuals to utilise based on the context. In this case, multiple modes of communicating a message are in question (i.e., text and images), both of which are indispensable for accomplishing the task outcomes. Therefore, it raises the question of whether the scope of focused tasks can be extended to involve a non-linguistic focus of tasks and be concerned about multimodality as well.

Making sound decisions regarding the dual category of here-and-now and there-and-then tasks is a bit problematic since the conceptualisations of these task types are based on few examples provided in the relevant literature. While the examples are mainly based on the presence or absence of the *task content* (visual or linguistic elements to be used when implementing a task) in Ellis's words (2020), the

concept needs further elaboration to clarify what it exactly denotes. For instance, task content can be a visual that learners need to implement a task or simply a visual that is not essential to be worked on while performing a task but can help contextualise and structure the task anyway so that learners can more easily perform it. This distinction affects whether a task is considered as a HN or TT task. Moreover, task content can be mistaken for task instructions by the professionals in the field to design, select or sequence the tasks, which also changes the decisions made as to the type of tasks to a great extent.

Another issue that makes the decision-making process regarding the task types is that a task can include a few steps, each of which bears different task type characteristics. To illustrate, even if the distinction between a closed and an open task may seem straightforward, one step of a task can have an open outcome as it allows learners to make choices. Still, their choices can be interpreted based on a limited number of criteria given in the next step as in closed tasks. Such activities were marked as OP since learners start with their own choices and the outcome changes based on their initial choice. Nonetheless, it can also be claimed that such a task also bears the features of CL tasks due to the limited number of solutions offered. In fact, it would not be surprising to observe such a situation in other tasks and with the other dual categories (e.g., HN versus TT). Therefore, considering tasks showing more or less of the characteristics of each task type in the dual categories as a continuum would provide a more accurate representation of their characteristics.

Task Complexity and Sequencing of Tasks

The last question of the present study aimed to explore how the task complexity dimensions of output-based and integrative tasks are sequenced in the EFL textbooks used in the Anatolian high schools in Turkey. To this end, the

frameworks by Robinson (2001a, 2001b, 2003, 2007, 2008, 2011, 2015) and Skehan and Foster (2001) were utilised.

First, the frequencies and percentages of tasks increased in task complexity along resource-directing and resource-dispersing dimensions were calculated and then, the complexity steps of the tasks based on SSARC model were determined. They consisted of Step 1 (-RDir -RDis), Step 2 (-RDir +RDis), a new, intermediary step generated (+RDir -RDis) and Step 3 (+RDir +RDis). These steps were considered to find out whether the tasks in the units follow a systematic progression from simple to complex as prescribed by the model.

Before discussing the task complexity dimensions detected in four EFL textbooks in the current study, it must be noted that only 66 OB and INT tasks were detected within 40 units in the textbooks. Although the frequencies and percentages of these tasks were calculated to be able to have a clear picture of the task complexity dimensions manipulated and sequenced, since a rather limited number of tasks were detected in each unit of each textbook, this should be taken into consideration when interpreting the study findings, especially the ones for the sequencing of tasks.

Among the resource-directing dimensions making cognitive/causal demands, *there-and-then* (TT) dimension was found to be the most manipulated dimension in grades 9, 10 and 11, whereas *causal reasoning* was the most manipulated one in grade 12. The common employment of TT tasks partially is in partial agreement with the findings of the study by Butler et al. (2018) as larger proportions of tasks made complex in this dimension were found in two or the four books in that study.

Except for the TT condition, less than 30 per cent of tasks were manipulated along the other resource-directing tasks in the textbooks. *Perspective-taking* (PT), for

example, was found to be rarely required in the tasks analysed in the current study. This finding is in good agreement with the previous findings (Butler et al., 2018). Besides, no tasks were found to be manipulated along *causal reasoning* in grade 9, *intentional reasoning* in grade 10 and *spatial reasoning* in grades 11 and 12. The low frequency of tasks that demand reasoning would lend support to the claim made by Lambert (2018) that the tasks that introduce communicative gaps based on reasoning are less in number when compared to the ones with information gap in L2 instructional materials. This finding also confirms the findings of the study by Butler et al. (2018) as the tasks requiring some kind of reasoning were found to be low in frequency in the EFL textbooks. Butler et al. (2018) interpret this finding as the insufficient consideration of reasoning demand.

Although the interpretation of Butler et al. (2018) seems reasonable, there are two issues to consider when interpreting the findings of the current study regarding the reasoning demands (CR, IR, SR). First, several tasks in the textbooks examined were not coded to necessitate reasoning since it was not entirely clear from the task instructions whether they are highly conducive to the reasoning types determined by Robinson (2001a, 2001b, 2003, 2007, 2008, 2011, 2015) as the verbs chosen for the instructions were sometimes equivocal in respect to the intended cognitive processes learners are to get engaged in. Therefore, the tasks were coded to require specific reasoning types only when the relevant expectations are explicitly stated through the task instructions.

The second issue to consider is that only a limited number of reasoning types are included in Robinson's (2001a, 2001b, 2003, 2007, 2008, 2011, 2015) framework. In Robinson's framework, reasoning is restricted to three types only (SR, CR, IR). The tasks in the textbooks that require some sort of reasoning were not

marked to necessitate reasoning when they do not fit the operationalisation of reasoning types in the framework in question. For this reason, the findings related to the reasoning demands in the present study should be interpreted with caution.

As to the resource-dispersing dimensions making performative/procedural demands, *single task* (ST) and *single step* (SS) dimensions were found to be the two most manipulated dimensions in each grade level. On the other hand, only a small proportion of the OB and INT tasks were manipulated along *task structure* and *prior knowledge* dimensions to increase complexity.

With respect to the -ST condition, more than half of the OB and INT tasks in the textbooks for grades 9, 11 and 12 necessitate performing more than one action concurrently. For instance, learners are expected to both speak and take notes or remember the details of an event and report it to their friends simultaneously on several tasks.

When it comes to the -SS condition, more than half of the OB and INT tasks in the textbooks for grades 9, 10 and 11 were detected to consist of at least one step. This is partially due to the fact that tasks require the transfer of information to fill a communicative gap. For instance, learners are required to (1) search for something on the internet, (2) save and (3) organise the data they have found and (4) share them with other learners.

As opposed to the ST and SS dimensions, the percentages of the tasks that are made complex along *task structure* (TS) and *prior knowledge* (PK) dimensions are very low. As for the TS dimension, it was noticed that learners are mostly supported in terms of both the expectations for the task steps to be followed and the contents of the task outcomes through models, templates and detailed questions provided. Regarding the PK dimension, it was discovered that learners are generally supported

to have enough background knowledge and successfully implement the tasks through the provision of topic-related discussion opportunities and relevant vocabulary.

Besides, pre-task activities within the same unit are geared towards familiarising learners with genre-related aspects and similar cognitive/conceptual demands.

It is worthy of noting that the textbooks are especially successful in the attempts made to introduce genre-related aspects through several activities which engage learners with such actions as identifying the genres, ordering text sections as well as determining and grouping the functions of text sections in relevant contexts.

After the OB and INT tasks were analysed for the resource-directing and resource-dispersing dimensions, the next step was to determine if the tasks in the same units are sequenced to increase complexity in line with the principles of SSARC by Robinson (2015). The OB and INT tasks were found to be either in the second (automatisation) or third (restructuring and complexity) stages of SSARC in most cases, the third stage (RC) stage being the most common one. While the tasks made complex only along resource-directing dimensions (+RDir) were the least common ones in the textbooks for grades 10, 11 and 12, the tasks that were at the first and the simplest stage of SSARC was the least common in grade level 9.

This finding, however, should not be interpreted as that learners are expected to perform tasks in their complex forms (RC stage) most of the time without being given any opportunities to get engaged in less demanding activities of similar sort previously. As a matter of fact, several activities that precede the tasks in the same units of the textbooks were excluded from the analysis for task complexity since they were found to fail some criteria of taskness. Therefore, it is not surprising that another finding to emerge from the data regarding the task complexity dimensions is

that no systematic progression from simple to complex tasks as prescribed in the SSARC model is detected in the EFL textbooks examined in the current study.

The findings related to the sequencing of the tasks based on task complexity dimensions are in line with the study by Butler et al. (2018), in which tasks in the EFL textbooks were not found to be manipulated and sequenced based on cognitive demands but based on linguistic elements.

Considering that no claim was made in the curriculum (MoNE, 2018) that the tasks were sequenced according to the task complexity dimensions based on TCF by Robinson (2011) or the SSARC model developed out of it, it would not be surprising to find that the task sequencing in the textbooks examined in the current study does not fit this model.

The textbooks for the secondary level are introduced to consist of “task-based activities” (MoNE, 2018, p. 6). Nevertheless, the curriculum guiding the textbooks is also claimed to be “multi-syllabus”, sharing components from different syllabus types, especially from “functional syllabus” and “skill-based syllabus”, and focusing on three areas of learning: *Functions and Useful Language*, *Language Skills and Learning Outcomes* as well as *Suggested Materials and Tasks* for each unit (MoNE, 2018, p.13). This would suggest that functions and skills were mainly considered in the syllabus design rather than the cognitive demands the activities make. This may also explain why sequencing was not done based on task complexity, which is associated with cognitive factors.

Implications for Practice

The first contribution of the study for practice might be that it provides material developers with an in-depth analysis of textbook contents as to the nature of non-tasks and tasks, task types and task complexity based on the studies by Ellis

(2003, 2001, 2012, 2020), Ellis and Shintani (2014), Robinson (2001a, 2001b, 2003, 2007, 2008, 2011, 2015) as well as Skehan and Foster (2001). The findings of the study, therefore, could be helpful for them in making appropriate decisions when designing and sequencing tasks for a specific learner group.

One of the most marked observations to emerge from the data was that the wording in the activity instructions may largely determine the focus of activity, turning into a task or a non-task (i.e., exercise). The design of activities in the analysed textbooks can be evaluated and modified in light of the theoretical framework for task characteristics if task-based activities are targeted in textbooks. For instance, it was previously noted that the activities in some sections such as *Culture*, *Pronunciation*, and *Let's Explore* do not qualify for tasks since they fail to meet the taskness criteria taken as a reference point. However, these sections could be redesigned to engage learners in meaningful tasks.

The activities in the *Culture* sections were discovered to present engaging and thought-provoking contents, which are unfortunately followed mainly by comprehension activities to check understanding. Such activities can be redesigned to have meaning-focus and present a communicative gap to give learners a real motivation for the accomplishment of a communicative outcome. The activities, for example, can instruct learners to share and compare their experiences, perspectives, cultures etc. for a project, solution of a problem or competition, which would also help design a variety of RW tasks.

It was previously mentioned that several discovery-oriented activities in the *Let's Explore* sections of the textbooks were coded non-task since they lack a communicative gap and do not allow learners to use their own (non) linguistic resources. It is also the case for the activities in the *Pronunciation* sections. When the

instructions of these activities are manipulated to provide learners with a communicative gap to be filled and allow learners to use their own (non) linguistic resources to work on language points and come up with a rule, they would be good candidates for focused tasks.

The professionals in the field of ELT can follow some strategies to transform *task-supported* textbooks into *task-based* ones. As the textbooks analysed in the present study do not take tasks as the core element for the whole syllabus, the first step would be to take tasks as the primary unit to determine and sequence the contents of a syllabus (Ellis et al., 2019; Révész, n.d.). For tasks to be the main components of a syllabus, non-task activities (i.e., exercises) can be adapted to qualify for a task. Ellis's (2009) and Ellis and Shintani's (2014) framework for the concept of task can be taken as a guideline to manipulate non-tasks.

It is argued that an exercise (i.e., non-task) may be made more task-like by changing its design features (Ellis, 2003), and previous research showed that task-like activities could be converted to tasks with a few simple modifications (Butler et al., 2018; Ellis, 2003; Harris, 2016; Rodríguez Sánchez, 2018; Terekhova, 2017; Willis, 2006). In a similar vein, material developers and language teachers can redesign some task-like activities bearing the potential to become tasks with minor changes in the workplan by utilising the information provided in the present study. In this way, the workplans that fall halfway between tasks and non-tasks (i.e., exercises) may turn into tasks during the implementation process.

To qualify for task, activities can be manipulated to focus on *meaning* rather than form (criterion 1). This can be realised if language points are presented or encouraged to be used in context, through which learners would focus on

contextualised meaning arising from the acts in communication. For this purpose, the activities can be created with situational or at least interactional authenticity.

A communicative *gap* can be created (criterion 2) in the activities so that learners feel the urge to communicate in L2. It can be done by introducing information to the learners in a split fashion (information gap) or asking them to share their preferences, feelings or attitudes on a debatable issue (opinion gap) or instructing them to infer from given information (reasoning gap). In this way, learners will not only practice L2 but also use it to transfer information and accomplish a goal. For instance, when learners prepare a poster, the instructions of the activity can be improved to give them a chance to share what they have produced with a relevant audience for a communicative goal.

Another way to transform non-tasks (exercises) into tasks would be to allow learners to depend on their own (*non*)*linguistic resources* instead of limiting them to the use of sample language explicitly presented to be performed (criterion 3). With this purpose in mind, sample linguistic features provided with the activities would be excluded. If the aim is to orient learners to specific linguistic features, then the task instructions can be carefully written to make these features very useful or highly likely to occur during the performance instead of explicitly asking their practice as in focused tasks.

Finally, learners can be given a specific outcome to be reached at the end of their performance (criterion 4). For instance, free conversation activities can be restructured to make learners arrive at a concrete outcome instead of merely asking them what they think. Some examples of outcomes can be voting for the best proposal, preparing a report to present to the whole class or arriving at a solution.

The present study could be helpful also for teachers to gain valuable insights into the concept of language task, which was found to be challenging to conceptualise and implement by language teachers (Dao & Newton, 2021; East, 2018; Erlam, 2016; Jaruteerapan, 2020; Oliver & Bogachenko, 2018; Zhu, 2020). It can assist language teachers in designing, utilising and adapting the tasks and task types based on their teaching context.

Since the present study only focused on tasks as workplans and did not deal with the task implementation process, all the activities were evaluated as they were presented in the textbooks. However, the workplan constitutes only one aspect of language activities. The way learners approach activities (Ellis, 2012) and the way the language teachers present an activity also play a role in whether an activity can be considered a task or non-task.

The present study revealed that task instructions have a critical role in determining both task outcomes and the cognitive processes learners are to be engaged in. First, the verbs chosen in the task instructions can turn a non-task (i.e., exercise) into a task with a communicative outcome. For instance, the verb *discuss* can mean many actions, such as *talk about*, *explain*, *compare*, *interpret* and *decide* to learners, some of which may only serve to practice L2, while some may lead learners to produce a communicative outcome such as an abstract idea. When it was unclear what was expected from the learners to arrive at the end of task performance, several activities were coded as non-task.

Even if some activities were coded as a task and found to have the potential to lead learners to perform certain cognitive processes determined in the task complexity dimensions, they were not coded to yield these cognitive processes as the verbs (e.g., *discuss*) in the task instructions were found to be too general or

ambiguous. When the instructions do not clarify what is intended for learners to perform and learn, they may promote fewer opportunities for meaningful L2 use to facilitate L2 development (Butler et al., 2018). Since it is well recognised that the verbs utilised in the textbooks can affect the cognitive complexity of tasks (Anderson et al., 2001), the verbs can be changed to pose the intended cognitive demands. Therefore, textbook designers can make use of Bloom's revised taxonomy (Anderson et al., 2001) to better communicate what cognitive processes learners are expected to be engaged in, to provide consistency and walk them through different levels of cognitive processes for their L2 development.

Based on the suggestions made above for task design, a sample task (Appendix) was developed by the researcher from a non-task activity, and it was examined in terms of task characteristics, task types and the task complexity dimensions. The sample task may help clarify how non-task activities can be manipulated to satisfy the task criteria by Ellis (2009) and Ellis and Shintani (2014). Besides, it may present a detailed examination of a task in terms of task characteristics, task types and task complexity dimensions. The sample is hoped to inform material developers and language teachers on how to design tasks and analyse task contents.

Implications for Further Research

To the best knowledge of the researcher, the present thesis is the first study that investigated the task characteristics, task types and task complexity dimensions within the ELT textbooks used in the Anatolian high school context in Turkey based on the theoretical frameworks by Ellis (2001, 2003, 2009, 2011, 2020) and Ellis and Shintani (2014), Robinson (2001a, 2001b, 2003, 2007, 2008, 2011, 2015) as well as Skehan and Foster (2001). One contribution of the present thesis to the field of ELT

would be that the findings inform the researchers in the field on how the concept of task, task types and task complexity are actualised in the EFL textbooks in Turkey. Nevertheless, some suggestions can be made for further research.

The first suggestion for future studies is that the scope of the current study can be broadened by including the other textbooks used in the Anatolian high school context that have not been written by the textbook writers in the Ministry of Education (MoNE). Besides, another study can be conducted by examining the other ELT textbooks that target a different learner profile, such as the learners that learn English for specific purposes or belong to different age groups.

The first stage of the data analysis required identifying the task characteristics the activities possess in the textbooks and coding them either as a task or a non-task (i.e., exercise) rather than taking the taskness as ranging along a continuum. With this purpose in mind, the task framework by Ellis (2009) and Ellis and Shintani (2014) was used as a point of reference. However, it was noticed that the activities in these textbooks marked as non-task based on the aforementioned taskness criteria generally satisfy two or three criteria out of four, but mostly fail to meet the fourth criterion related to having a communicative *outcome*. Although these activities were considered non-task and excluded from the second and third stages of the analysis due to the requirements of the present study, it would not be wrong to note that they still can be considered *task-like*. Although some activities are tentatively called *task-like* in the present study to indicate that they possess some of the task characteristics defined, there was not used a definite category to classify activities as *task-like* when analysing activities since there is no specific, widely accepted definition proposed to make it clear what the concept of *task-like* involves. Future studies that focus on

taskness as a continuum can offer a definition of task-likeness and discuss what particular features it would involve.

Another recommendation would be to replicate this study by employing a different research design. In the present study, only content analysis based on four ELT textbooks was conducted. For this very reason, it would be worthwhile to further explore task demands in the textbooks through triangulation by integrating semi-structured interviews with the textbook writers to also gain insight into their own understanding of task, task types and task complexity.

In the present study, cognitive complexity dimensions within Robinson's framework (2001a, 2001b, 2003, 2007, 2008, 2011, 2015) were primarily utilised to analyse task complexity. While exploring the tasks in the textbooks, it was noticed that several tasks were marked to belong to the first and the simplest stage (stable attractor) of the SSARC model as they were not manipulated along either research-directing or research-dispersing (-RDir -RDis) dimensions. Nevertheless, it would be wrong to interpret the findings of the present study as several of the tasks in the textbooks examined do not pose any cognitive complexity at all. Since they were analysed primarily based on Robinson's framework (2001a, 2001b, 2003, 2007, 2008, 2011, 2015), the cognitive processes that are not included in this framework were not considered. This task complexity framework is mainly for task sequencing; nevertheless, it seems that several cognitive processes are not considered for the demands they would make on learners. Future research may examine the cognitive demands of the tasks from different lenses drawing from other task complexity frameworks in addition to the one by Robinson (2001a, 2001b, 2003, 2007, 2008, 2011, 2015).

It should be noted that the findings of this study are limited to the activities as workplan in the EFL textbooks and do not present any data regarding either the task-as-process or task outcomes. It means that the roles of the context (e.g., setting, time, purpose) and stakeholders (e.g., teachers, students) have not been adequately considered during the task implementation process. However, not only how a task is designed but also how it is performed affects the kind of language task-takers produce (Ellis, 2012). Hence any research into the phases of task-as-workplan and task-as-process together can give an insight into the connection or discrepancy between these phases. For instance, a micro-evaluation of tasks can be carried out to check if the activity during task-as-process occurs as predicted in the task-as-workplan (Ellis, 2018).

As for task complexity, only the cognitivist perspective was adopted in the present study. However, further studies could be carried out with teachers and students as to their perception of task demands drawn from *Sociocultural Theory* to compare and contrast the hypotheses of *Cognitive Hypothesis* and the findings of the previous studies. In this way, the demands tasks would pose on learners would be considered not only from a theory-driven, etic perspective but also from a participant-relevant, emic perspective, and holistic analysis, and evaluation of tasks would be made.

In addition to the tasks prepared for EFL textbooks, the design and implementation of digital tasks should also be examined as the L2 learning context is not limited to in-class activities anymore. Digital tasks may differently affect the complexity of tasks and introduce new demands, thus necessitating the redefinition of some dimensions within the TCF by Robinson (2011) or the generation of a new framework from scratch that would specifically pertain to digital tasks to better

reflect the nature of the tasks designed to be implemented in the digital sphere. Such studies would be helpful also in informing task designers when developing tasks for the digital sphere (e.g., language management systems) for English language teaching.

It is worth noting that the differences between the textbooks in terms of the proportion of the tasks, as well as the sequencing of the task types and task complexity levels revealed in the present study should not be taken to make claims about their relative qualities. The overall aim of this study was to examine tasks in those textbooks and compare and contrast them based on the issues above, not to make quality judgments. Further analysis will be needed if such an evaluation of the textbooks is aimed to be conducted.

The recommended studies aforementioned could be of great value in reconciling the understandings of different stakeholders (textbook designers, pre-service and in-service EFL teachers, L2 learners etc.) involved in the English language teaching and learning process of tasks.

Limitations and Delimitations

The results of the current study presented some information about tasks, task types and task complexity dimensions in the EFL textbooks used in the Anatolian high school context in Turkey. Nonetheless, the limitations and delimitations of the present study should be noted for further research.

The first delimitation lends itself to the research design employed. Since the study was based on the content analysis of the tasks as workplans (i.e., design material), it disregarded task-as-process (i.e., the context of the task implementation), which can influence the task demands to a great extent. As the data was not triangulated through additional data sources and instruments such as classroom

observation for task implementation and semi-structured interviews with textbook writers, language teachers, or language learners, the findings are limited solely to the task-as-workplan in the textbooks analysed.

The second delimitation of this study might have its source in the sampling method of the textbooks analysed. The sample of the study was determined through purposeful sampling to gain insights into the language tasks exclusively used in the Anatolian high schools in Turkey. However, there are alternative textbooks prepared by commercial publishing houses and recognised by the MoNE to be covered in the same context. Moreover, there are different types of high schools in Turkey that provide their students with different kinds of textbooks than the ones examined within the scope of this study. Therefore, the findings of the study cannot be generalised to the conditions of the tasks as workplan in the textbooks used neither in all Anatolian High schools nor all types of high schools in Turkey.

Another delimitation of the present study is related to the sections in the textbooks that were excluded from the study. The *Check Yourself* sections in the textbooks for grades 9 and 11 were not included in the analysis as they are included only to help learners reflect on their learning, performance and progress (MoNE, 2018). Besides, the *Grammar Reference* section at the end of the textbook for 11th graders was not included in the analysis since it was designed as a reference section for the activities within the textbook. In addition, the questions on the unit cover pages as well as the *Game* sections within the textbooks for grades 9 and 11 were not analysed in the present study as they lack any instructions provided to make it clear what is expected from learners.

As for the delimitation regarding the task phases, the distinction between the pre-task, main task (during task) or post-task phases was disregarded within the

scope of the present study for three reasons. First, even if the EFL textbooks examined in the present study are claimed to consist of task-based activities (MoNE, 2018), they do not necessarily follow the three phases of the task-based language teaching framework as ideally prescribed by Willis (1996) or Ellis (2006a). The curriculum for high schools in Turkey is not task-based and shares components from different syllabus types, especially from “functional syllabus” and “skill-based syllabus” (MoNE, 2018, p. 13). Second, although all the phases are functional to enable learners to improve their language, Ellis (2018) maintains that only the performance of the main task in the *during task* phase is obligatory, whereas the others are optional, thus indicating that a task may lack pre-and/or post-task phases. In addition, the present study does not necessitate the classification of task phases by any means. Hence, any activity that meets the taskness criteria by Ellis (2009) and Ellis and Shintani (2014) was considered a task within the scope of this study and no further attempt was made to specify the task phases.

Another delimitation worthy of mentioning is that the scope of the present study is solely limited to the frameworks by Ellis (2001, 2003, 2009, 2011, 2012, 200), Ellis and Shintani (2014), Robinson (2001a, 2001b, 2003, 2007, 2008, 2011, 2015) as well as Skehan and Foster (2001) used for the operationalisations of the tasks, task types and task complexity, respectively. Hence, the decisions made concerning the *taskness* of the activities, task types and task complexity dimensions were bound to these frameworks. Nevertheless, quite different results would have been arrived at if different frameworks had been drawn on for analysing the tasks. Therefore, the results and the discussions of this study should be evaluated accordingly.

The final delimitation of the study is concerned with the cognitive complexity dimensions examined in the present study. Although Checklist 3 draws from TCF by Robinson (2001a, 2001b, 2003, 2007, 2008, 2011, 2015), three dimensions from that framework, *few elements* and *independency of steps* and *planning time*, were excluded from the scope of the study. *Few elements* dimension was excluded as its operationalisation necessitates the comparison of two tasks with different numbers of elements, which was not possible to operationalise for a study examining tasks in the textbooks (Barrot, 2016; Butler et al., 2018).

Secondly, *independency of steps* dimension was excluded since all the tasks with two or more subsequent steps in the present study were composed of a sequence of activities that makes the final activity to depend on the previous ones for the successful accomplishment of the task outcome.

Finally, *planning time* dimension was not examined in this study. The instructions in the textbooks do not make it possible to determine whether task takers are provided with either *pre-task planning* (also called *strategic planning*) or online planning time as manipulated in previous studies (Ellis & Yuan, 2004; Tabari, 2018; Yuan, 2001; Yuan & Ellis, 2003). Besides, even if *task repetition* was taken within the scope of *planning time* in some studies (Ahmadian & Tavakoli, 2010; Khatib & Farahanynia, 2020), it is associated with *prior knowledge* dimension in the present study as the repetition of a task would also familiarise task-takers with the cognitive/conceptual demands, genre-specific restrictions or the topic of that task.

As for the limitations of the study, the first issue that needs to be discussed is related to the decisions made regarding the taskness of the activities. One factor that makes the task analysis challenging is that tasks can be unclear and consist of hidden variables (Ellis et al., 2019, p. 96).

It was also aforementioned that activities may bear the characteristics of both tasks and non-tasks (i.e., exercises), and there may not be a watertight distinction between them all the time. Therefore, the *taskness* of the activities is often suggested to be taken as a continuum rather than a dichotomy in the English language teaching literature.

For the most part, it was rather challenging to decide whether an activity satisfies a criterion in its fullest sense in the present study. Making decisions regarding criterion 3 (i.e., the use of (non)linguistic resources), especially in input-based activities constitutes a good example of this as learners are afforded the (non)linguistic content, leaving them with limited opportunities to visit their own resources. However, a distinction was made between tasks and non-tasks in the present study for the study had to focus exclusively on tasks to answer the second and the third questions of the present study.

It should be noted that solely output-and integrative tasks were investigated for task complexity dimensions in this study as the task complexity framework proposed by Robinson (2001a, 2001b, 2003, 2007, 2008, 2011, 2015), which was used as a primary framework for task complexity dimensions, does not address the demands posed by the task content or input and makes predictions on the effect of task complexity on L2 production only. Therefore, task complexity dimensions of input-based tasks (i.e., listening and reading) were not addressed in this study, thus constituting the second limitation of the study.

Finally, the limitation due to the lack of clear task structure is worth noting. A task is expected to have a clear structure that makes its purpose, application, and the desired outcome clear. However, several tasks in the textbooks examined were found to lack a clear structure to enable the researcher and the coders to determine exactly

what task components or steps are involved, thus making it hard to arrive at decisions regarding not only the task characteristics but also the types and complexity dimensions of tasks. This was the case in a previous study as well, where the “insufficiency of structured input in their representation” in an EFL textbook made it difficult to determine whether an activity qualify for task (Günay, 2007, p. 117). In such cases, as Nunan (1989, p. 11) also admits, “making decisions [was] partly intuitive and judgemental”.

All in all, the current research was an attempt to provide a comprehensive analysis on language tasks in the ELT textbooks used in the Anatolian high school context in Turkey. Specifically, it aimed to shed light on the task characteristics, task types, and task complexity dimensions in these textbooks. Despite the limitations mentioned above, it is hoped that the findings of the study will support the stakeholders in the relevant field to better conceptualise the notions of task, task type and task complexity and make informed decisions for the design and implementation of tasks in English language teaching. The findings of the study are also expected to contribute to the research on task design and task-based syllabus design in the field of ELT.

REFERENCES

- Adams, R., & Alwi, N. A. N. (2014). Prior knowledge and second language task production in text chat. In M. González-Lloret & L. Ortega (Eds.), *Technology-mediated TBLT: Researching technology and tasks* (pp. 51-78). John Benjamins.
- Ahmadian, M. J., & Tavakoli, M. (2010). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research*, 15(1), 35-59.
<https://doi.org/10.1177/1362168810383329>
- Ahour, T., & Shemshadsara, Z. G. (2015). The impact of focused and unfocused tasks on university students' grammatical achievement. *Theory and Practice in Language Studies*, 5(1), 124-133. <http://dx.doi.org/10.17507/tpls.0501.17>
- Akbulut, N. (2014). *Teaching English to young learners through task-based language teaching* [Unpublished master's thesis]. Pamukkale University.
- Akdağ, E., Baydar Ertopçu, F., Uyanık Bektaş, K., Umur Özadalı, S., & Kaya, T. (2019). *Silver lining 11 (Student's book)*. Milli Eğitim Bakanlığı.
- Akın, A. (2020). *Students' attitudes towards and beliefs about task based language teaching and its effects on language self-efficacy beliefs* [Unpublished master's thesis]. Ufuk University.
- Alavinia, P., Shafaei, A., & Salimi, A. (2018). The effects of focused/unfocused audio-appended reading tasks on intermediate female EFL learners' written accuracy. *International Journal of Instruction*, 11(2), 203-220.
- Alemi, M., & Hesami, Z. (2014). The presentation of different types of tasks in ELT textbooks. *Teaching English Language*, 8(1), 61-92.

- Alemi, M., Jahangard, A., & Hesami, Z. (2013). A comparison of two global ELT course books in terms of their task types. *IARTEM e-Journal*, 5(2), 42-63.
<https://doi.org/10.21344/iartem.v5i2.767>
- Alwi, N. A. N. M. (2010). *Examining the language learning potential of a task-based approach to synchronous computer-mediated communication* [Unpublished doctoral dissertation]. Victoria University of Wellington.
- Alwi, N. A. N. M., Adams, R., & Newton, J. (2012). Writing to learn via text chat: Task implementation and focus on form. *Journal of Second Language Writing*, 21(1), 23–39. <https://doi.org/10.1016/j.jslw.2011.12.001>
- Anderson, L. W., Krathwohl, D. R., Airasian, E. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Arslanyılmaz, A., & Pedersen, S. (2010). Enhancing negotiation of meaning through task familiarity using subtitled videos in an online TBLL environment. *The Turkish Online Journal of Educational Technology*, 9(2), 64–77.
- Asma, B. (2018). Examining the role of task-based language teaching in fostering EFL learners' attitudes and motivation. *Journal of Education and New Approaches* 1(1), 53-60.
- Awwad, A. A. (2017). *The effects of task complexity manipulated by intentional reasoning demands on second language learners' speech performance: interaction with language proficiency and working memory* [Unpublished doctoral dissertation]. University of Reading.

- Awwad, A., Tavakoli, P., & Wright, C. (2017). “I think that's what he's doing”: Effects of intentional reasoning on second language (L2) speech performance. *System*, 67, 158-169.
- Baklacı, Ö. (2018). *Yabancı dil olarak Türkçe öğretiminde görev odaklı yöntem ışığında yazma becerisinin geliştirilmesi* [Unpublished master's thesis]. İstanbul University.
- Baralt, M. (2010). *Task complexity, the cognition hypothesis, and interaction in CMC and FTF environments* [Unpublished doctoral dissertation]. Georgetown University.
- Baralt, M., Gilabert, R., & Robinson, P. (2014). An introduction to theory and research in task sequencing and instructed second language learning. In M. Baralt, R. Gilabert, & P. Robinson (Eds.), *Task sequencing and instructed second language learning* (pp. 1-34). Bloomsbury.
- Barrot, J. S. (2019). Examining the task complexity in ELT coursebooks. *The Asia-Pacific Education Researcher*, 28(6), 469-481. <https://doi.org/10.1007/s40299-019-00448-2>
- Berben, M., Van den Branden, K., & Van Gorp, K. (2007). “We’ll see what happens”: Tasks on paper and tasks in a multilingual classroom. In K. Van den Branden, K. Van Gorp, & M. Verhelst (Eds.), *Tasks in action: Task-based language education from a classroom-based perspective* (pp. 32–67). Cambridge Scholars Publishing.
- Bistline Bonilla, C. (2020). *Interaction in synchronous computer-mediated communication: The effects of interlocutor, task, and state anxiety* [Unpublished doctoral dissertation]. Georgetown University.

- Breen, M. (1989). The evaluation cycle for language learning tasks. In R. K. Johnson (Ed.), *The second language curriculum* (pp. 187–206). Cambridge University Press.
- Brown, J. D., Hudson, T. D., Norris, J. M., & Bonk, W. J. (2002). *An investigation of second language task-based performance assessments*. University of Hawai'i Press.
- Bulut, E., Baydar Ertopçu, F., Umur Özadali, S., & Şentürk, S. (2019). *Teenwise 9 (Student's book)*. Milli Eğitim Bakanlığı.
- Butler, Y. G., Kang, K. I., Kim, H., & Liu, Y. (2018). 'Tasks' appearing in primary school textbooks. *ELT Journal*, 72(3), 285-295.
<https://doi.org/10.1093/elt/ccx056>
- Byram, M., Hu, A. (2013). *Routledge encyclopedia of language teaching and learning*. Routledge.
- Çelik, D. (2021). *Türkçenin yabancı dil olarak öğretiminde A2 seviyesi konuşma etkinliklerinin görev odaklı yaklaşıma göre incelenmesi* [Unpublished master's thesis]. Yıldız Technical University.
- Ceylan, T. (2016). *The effectiveness of task-based instruction in improving students' writing and speaking skills* [Unpublished master's thesis]. Ufuk University.
- Cho, H. (2011). *Negotiation of meaning in synchronous computer-mediated communication in relation to task types* [Unpublished doctoral dissertation]. Indiana University Bloomington.
- Choong, K. P. (2011). Task complexity and linguistic complexity: An exploratory study. *Studies in Applied Linguistics and TESOL*, 11(1), 1-28.
<https://doi.org/10.7916/salt.v11i1.1413>

- Choong, K. P. (2014). *Effects of task complexity on written production in L2 English* [Unpublished doctoral dissertation]. Columbia University.
- Çimen, F., Taşkiran Tiğın, B., Çokçalışkan, A., Özyıldırım, N., & Özdemir, M. (2019a). *Count me in 12 (Student's book)*. Milli Eğitim Bakanlığı.
- Çimen, F., Taşkiran Tiğın, B., Çokçalışkan, A., Özyıldırım, N., Özdemir, M., Demir, E., & Fındıkçı, G. (2019b). *Count me in 10 (Student's book)*. Milli Eğitim Bakanlığı.
- Cobb, M. (2010). *Meta-analysis of the effectiveness of task-based interaction in form-focused instruction of adult learners in foreign and second language teaching* [Unpublished doctoral dissertation]. University of San Francisco.
- Cobb, M., & Lovick, N. (2007, September 20-22). *The concept of foreign language task, misconceptions and benefits in implementing task-based instruction* [Paper presentation]. 2nd International Conference on Task-Based Language Teaching, Honolulu, HI, United States.
- Coughlan, P., & Duff, P. A. (1994). Same task, different activities: Analysis of SLA task from an activity theory perspective. In J. Lantolf, & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 173-193). Ablex Publishing Corporation.
- Cunningham, S., & Moor, P. (2005). *New cutting edge: Elementary (Student's book)*. Longman.
- Dao, H., & Newton, J. (2021). TBLT perspectives on teaching from an EFL textbook at a Vietnam university. *Canadian Journal of Applied Linguistics*, 24(2), 99-126. <https://doi.org/10.37213/cjal.2021.31371>
- De Swert, K. (2012). Calculating inter-coder reliability in media content analysis using Krippendorff's Alpha. *Center for Politics and Communication*, 15.

- Değirmenci Mutlu, S. (2016). *A case study: The potential impact of task-based instruction on EFL learners' writing self-efficacy* [Unpublished master's thesis]. Bahçeşehir University.
- Donate, Á (2018). *Cognitive task complexity, foreign language anxiety and L2 performance in Spanish: A task-based language teaching perspective* [Unpublished doctoral dissertation]. Georgetown University.
- Drisko, J. W., & Maschi, T. (2016). *Content analysis*. Pocket Guides to Social Work Research Methods.
- East, M. (2018). How do beginning teachers conceptualise and enact tasks in school foreign language classrooms?. In V. Samuda, K. Van den Branden & M. Bygate (Eds.), *TBLT as a researched pedagogy* (pp. 23-50). John Benjamins.
<https://doi.org/10.1075/tblt.12>
- Educational Informatics Network (EBA). (2021, March 21).
<https://www.eba.gov.tr/arama?q=%C4%B0ngilizce%20kitap>
- Ellis, R. (1995). Interpretation tasks for grammar teaching. *TESOL Quarterly*, 29(1), 87-105. <https://doi.org/10.2307/3587806>
- Ellis, R. (2000). Task-based research and language pedagogy. *Language Teaching Research*, 4(3), 193-220. <https://doi.org/10.1177/136216880000400302>
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- Ellis, R. (2005). Principles of instructed language learning. *System*, 33(2), 209-224.
<https://doi.org/10.1016/j.system.2004.12.006>
- Ellis, R. (2006). The methodology of task-based teaching. *Asian EFL Journal*, 8(3), 19-45.

- Ellis, R. (2008). Evaluating and researching grammar consciousness raising tasks. In P. R. Dickens & K. Germaine (Eds.), *Managing evaluation and innovation in language teaching* (pp. 220–52). Longman.
- Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19(3), 221–246.
<https://doi.org/10.1111/j.1473-4192.2009.00231.x>
- Ellis, R. (2011). Macro- and micro-evaluations of task-based teaching. In B., Tomlinson (Ed.), *Materials development in language teaching* (pp. 212–235). Cambridge University Press.
- Ellis, R. (2012). *Language teaching research and language pedagogy*. Wiley-Blackwell.
- Ellis, R. (2013). *Task-based language teaching: Responding to the critics*. University of Sydney Papers in TESOL, 8.
- Ellis, R. (2014, December 4–6). *Taking the critics to task: The case for task-based teaching* [Conference Session]. CLaSIC, National University of Singapore, Singapore.
- Ellis, R. (2015). *Understanding second language acquisition* (2nd ed.). Oxford University Press.
- Ellis, R. (2016). Focus on form: A critical review. *Language Teaching Research*, 20(3), 405–428. <https://doi.org/10.1177/1362168816628627>
- Ellis, R. (2017). Moving task-based language teaching forward. *Language Teaching*, 50(4), 507–526. <https://doi.org/10.1017/S0261444817000179>
- Ellis, R. (2018). *Reflections on task-based language teaching*. Multilingual Matters.
- Ellis, R. (2020, September 8). *Using tasks in language teaching* [Webinar]. Cambridge Live Experience.

https://www.youtube.com/watch?v=jsBTQgE8uhw&ab_channel=CambridgeUniversityPressELT

- Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. Routledge.
- Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, 26(1), 59-84. <https://doi.org/10.1017/S0272263104026130>
- Ellis, R., Skehan, P., Li, S., Shintani, N., & Lambert, C. (2019). *Task-based language teaching: Theory and practice*. Cambridge University Press.
- Erlam, R. (2016). 'I'm still not sure what a task is': Teachers designing language tasks. *Language Teaching Research*, 20(3), 279-299. <https://doi.org/10.1177/1362168814566087>
- Erlam, R., & Ellis, R. (2018). Task-based language teaching for beginner-level learners of L2 French: An exploratory study. *Canadian Modern Language Review*, 74(1), 1-26.
- Estaire, S., & Zanon, J. (1994). *Planning classwork: A task-based approach*. Macmillan Heinemann.
- Fraenkel, J. R., & Wallen, N. E. (2009). *Online learning center to accompany how to design and evaluate research in education* (7th ed.). McGraw-Hill Education.
- Frear, M. W. (2013). *The effects of cognitive task complexity on writing* [Unpublished doctoral dissertation]. Auckland University of Technology.
- Gilabert, R. (2007). The simultaneous manipulation of task complexity along planning time and (+/? Here-and-Now): Effects on L2 oral production. In M. P. Garcia Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 44–68). Multilingual Matters.

- Gilabert, R., Barón, J., & Llanes, À. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during oral performance. *IRAL - International Review of Applied Linguistics in Language Teaching*, 47(3-4), 367-395. <https://doi.org/10.1515/iral.2009.016>
- Gilabert, R., Manchón, R., & Vasylets, O. (2016). Mode in theoretical and empirical TBLT research: Advancing research agendas. *Annual Review of Applied Linguistics*, 36, 117-135. <https://doi.org/10.1017/S0267190515000112>
- Gökgöz, N. (2009). *The effects of task complexity on measures of accuracy and lexical variety in EFL writing* [Unpublished master's thesis]. Marmara University.
- Gonzalez-Lloret, M., & Ortega L. (2014). Towards technology-mediated TBLT: An introduction. In M. González-Lloret, & L. Ortega (Eds.), *Technology-mediated TBLT: Researching technology and tasks* (pp. 1-22). John Benjamins.
- Günel Şahan, S. (2019). *Görev temelli dil öğretim yönteminin öğrencilerin İngilizce dersine yönelik akademik başarılarına ve tutumlarına etkisi* [Unpublished master's thesis]. Akdeniz University.
- Günay, D. (2007). *The exploration of tasks in the 4th grade ELT coursebook used in state primary schools in Turkey* [Unpublished master's thesis]. İstanbul University.
- Halıcı Page, M. (2016). *A comparative study of task-based and traditional instruction on student motivation and vocabulary development in 7th grade Turkish EFL classrooms* [Unpublished master's thesis]. Bahçeşehir University.
- Hamavandy, M., & Golshan, M. (2015). Differential potential of SLA output tasks versus input-based teaching of English grammar: A comparative study. *Theory and Practice in Language Studies*, 5(10), 2083-2090. <http://dx.doi.org/10.17507/tpls.0510.15>

- Hardy, I. M., & Moore, J. L. (2004). Foreign language students' conversational negotiations in different task environments. *Applied Linguistics*, 25(3), 340-370.
<https://doi.org/10.1093/applin/25.3.340>
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), 77-89.
- Hsu, H. C. (2012). Investigating the effects of planning on L2 text-chat performance. *CALICO Journal*, 29(4), 619-638.
- Hsu, H. C. (2020). The impact of task complexity on patterns of interaction during web-based asynchronous collaborative writing tasks. *System*, 93, 1-13.
<https://doi.org/10.1016/j.system.2020.102328>
- Huda, N., & Syaifei, A. F. R. (2020). Task types found in the textbooks “Bahasa Inggris: When English rings a bell” and BRIGHT an English course for junior high school. *Journal of English Language Teaching*, 9(1), 210-218.
<https://doi.org/10.24036/jelt.v9i1.107932>
- İhtiyaryer, B. (2018). *Türkçenin yabancı dil olarak öğretiminde görev odaklı yaklaşım açısından kelime öğretimi* [Unpublished master's thesis]. İstanbul University.
- Ishikawa, T. (2006). The effect of manipulating task complexity along the (±here-and-now) dimension on L2 written narrative discourse. In M. D. P. G. Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 136-156). Multilingual Matters.
- Ishikawa, T. (2008). The effect of task demands of intentional reasoning on L2 speech performance. *Journal of Asia TEFL*, 5(1), 29-63.
- Iwashita, N., McNamara, T., & Elder, C. (2001). Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information-processing approach

- to task design. *Language Learning*, 51(3), 401-436. <https://doi.org/10.1111/0023-8333.00160>
- Jackson, D. O., & Suethanapornkul, S. (2013). The cognition hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning*, 63(2), 330-367. <https://doi.org/10.1111/lang.12008>
- Jaruteerapan, P. (2020). *The emerging understandings and practices of task-based language teaching (TBLT) by Thai EFL student teachers* [Unpublished doctoral dissertation]. Victoria University of Wellington.
- Jaruteerapan, P. (2021). A review of issues about teachers' understanding of the construct of tasks in task-based language teaching. *Journal of Liberal Arts Prince of Songkla University*, 13(2), 1-19.
- Kamalı, T. (2012). *Students' experiences and perceptions of anxiety, motivation, and self-confidence in speaking English during task-based language learning activities in second life: The case of METU* [Unpublished master's thesis]. Middle East Technical University.
- Kang, S., & Lee, J. H. (2019). Are two heads always better than one? The effects of collaborative planning on L2 writing in relation to task complexity. *Journal of Second Language Writing*, 45, 61-72. <https://doi.org/10.1016/j.jslw.2019.08.001>
- Karakoç, B. (2016). *Otantik görev odaklı uygulamaların yabancı dil eğitime etkisinin incelenmesi* [Unpublished doctoral dissertation]. Gaziantep University.
- Kayır, G. (2018). *Türkçe ve Türk Kültürü Dersi'nde görev temelli etkinliklerin İsviçre'de yaşayan Türk çocuklarının akademik başarı ve motivasyonlarına etkisi* [Unpublished doctoral dissertation]. Eskişehir Osmangazi University.
- Khatib, M. & Farahanynia, M. (2020). Planning conditions (strategic planning, task repetition, and joint planning), cognitive task complexity, and task type: Effects

on L2 oral performance. *System*, 93, 1-12.

<https://doi.org/10.1016/j.system.2020.102297>

Kim, Y. (2009). *The role of task complexity and pair grouping on the occurrence of learning opportunities and l2 development* [Unpublished doctoral dissertation]. Northern Arizona University.

Kırtaş, G. (2016). *ELT teachers' perceptions and practices of task-based language teaching* [Unpublished master's thesis]. Muğla Sıtkı Koçman University.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Addison-Wesley Longman Limited.

Krippendorff, K. (2011). *Computing Krippendorff's alpha-reliability*. Departmental Paper.

https://repository.upenn.edu/cgi/viewcontent.cgi?article=1043&context=asc_papers

Kuiken, F., & Vedder, I. (2007). Task complexity and measures of linguistic performance in L2 writing. *De Gruyter Mouton*, 45(3), 261-284.

<https://doi.org/10.1515/iral.2007.012>

Lai, C., & Li, G. (2011). Technology and task-based language teaching: A critical review. *CALICO Journal*, 28(2), 498-521.

Lambert, C. (2018). Tasks versus exercises. In J. I., Liantas & TESOL International Association (Eds.), *The TESOL encyclopedia of English language teaching* (pp. 1-6). John Wiley & Son.

Lambert, C. (2019). Task-based syllabus design. In R., Ellis, P., Skehan, S., Li, N., Shintani, & C., Lambert (Eds.), *Task-based language teaching: Theory and practice* (pp. 179-207). Cambridge University Press.

- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.
- Larson, G. M. (2020). *Paired-speaking tasks: planning conditions and task performance* [Unpublished doctoral dissertation]. Northern Arizona University.
- Lee, J. (2018). *The interactive effects of task complexity, task condition, and cognitive individual differences on L2 writing* [Unpublished doctoral dissertation]. University of Maryland. <https://doi.org/10.13016/hgvz-qwnr>
- Lee, J. (2020). Task closure and task complexity effects on L2 written performance. *Journal of Second Language Writing*, 50, 1-13.
<https://doi.org/10.1016/j.jslw.2020.100777>
- Levkina, M., & Gilabert, R. (2014). Task sequencing in the L2 development of spatial expressions. In M. Baralt, R. Gilabert, & P. Robinson (Eds.), *Task sequencing and instructed second language learning* (pp. 47-81). Bloomsbury.
- Littlewood, W. (2004). The task-based approach: Some questions and suggestions. *ELT Journal*, 58(4), 319-326.
- Long, M. H. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 77-99). Multilingual Matters.
- Long, M. H. (2015). *Second language acquisition and task-based language teaching*. John Wiley & Sons.
- Long, M. H. (2016). In defense of tasks and TBLT: Nonissues and real issues. *Annual Review of Applied Linguistics*, 36, 5-33.
- Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes, & S. Gass (Eds.), *Tasks and language learning* (pp. 123-167). Multilingual Matters.

- Luo, S. (2007). *Re-examining factors that affect task difficulty in TBLA* [Unpublished doctoral dissertation]. The Chinese University of Hong Kong.
- Malicka, A. (2014a). *The role of task complexity and task sequencing in L2 monologic oral production* [Unpublished doctoral dissertation]. Universitat de Barcelona.
- Malicka, A. (2014b). The role of task sequencing in monologic oral production. In M. Baralt, R. Gilabert, & P. Robinson (Eds.), *Task sequencing and instructed second language learning* (pp. 1-34). Bloomsbury.
- Malicka, A. (2020). The role of task sequencing in fluency, accuracy, and complexity: Investigating the SSARC model of pedagogic task sequencing. *Language Teaching Research*, 24(5), 642-665. <https://doi.org/10.1177/1362168818813668>
- Mehmood, H. (2021). *Exploring EFL instructors' perceptions of task-based language teaching* [Unpublished master's thesis]. İstanbul Aydın University.
- Michel, M. (2011). *Cognitive and interactive aspects of task-based performance in Dutch as a second language* [Unpublished doctoral dissertation]. University of Amsterdam.
- Minas, E. C. (2017). *Complexity in task-based language teaching and learning of isiXhosa as a second language in primary school* [Unpublished doctoral dissertation]. Stellenbosch University.
- Ministry of National Education (MoNE). (2018a). *İngilizce dersi öğretim programı (İlkokul ve ortaokul 2, 3, 4, 5, 6, 7 ve 8. sınıflar)*. Milli Eğitim Bakanlığı.
- Ministry of National Education (MoNE). (2018b). *Ortaöğretim İngilizce dersi (9., 10., 11. ve 12. sınıflar) İngilizce öğretim planı*. Milli Eğitim Bakanlığı.
- Ministry of National Education (MoNE). (2021, March 21). *EBA*. <https://www.eba.gov.tr/arama?q=%C4%B0ngilizce%20kitap>

- Moradi, A., & Farvardin, M. T. (2019). Negotiation of meaning by mixed- proficiency dyads in face- to- face and synchronous computer-mediated communication. *TESOL Journal*, 11(1), 1-17. <https://doi.org/10.1002/tesj.446>
- Motteram, G., & Thomas, M. (2010). Afterword: Future directions for technology-mediated tasks. In M. Thomas & H. Reinders (Eds.), *Task-based language learning and teaching with technology* (pp. 218-238). Continuum.
- Newton, J. (2001). Options for vocabulary learning through communication tasks. *ELT Journal*, 55(1), 30–37. <https://doi.org/10.1093/elt/55.1.30>
- Newton, J., & Bui, T. (2017). 12. Teaching with tasks in primary school EFL classrooms in Vietnam. In M. Ahmadian & M. García Mayo (Ed.), *Recent perspectives on task-based language learning and teaching* (pp. 259-278). De Gruyter Mouton. <https://doi.org/10.1515/9781501503399-013>
- Nitta, R., & Gardner, S. (2005). Consciousness-raising and practice in ELT coursebooks. *ELT Journal*, 59(1), 3-13. <https://doi.org/10.1093/elt/cci001>
- Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal*, 47(3), 203-210.
- Norris, J. M., Brown, J. D., Hudson, T. D., & Yoshioka, J. K. (1998). *Designing second language performance assessments*. University of Hawai'i Press.
- Nuevo, A. M. (2006). *Task complexity and interaction: L2 learning opportunities and development* [Unpublished doctoral dissertation]. Georgetown University.
- Nunan, D. (1988). *Syllabus design*. Oxford University Press.
- Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge University Press.
- Nunan, D. (2004). *Task based language teaching*. Cambridge University Press.

- Oliver, R., & Bogachenko, T. (2018). Teacher perceptions and use of tasks in school. In Samuda, V., Van den Branden, K., & M. Bygate (Eds.), *TBLT as a researched pedagogy* (pp. 72–95). John Benjamins.
- Oxford, R. L. (2017). Conditions for second language (L2) learning. In S. May (Ed.), *Encyclopedia of language and education* (3rd ed., pp. 27-41). Springer.
- Pallotti, G. (2019). Assessing tasks: The case of interactional difficulty. *Applied Linguistics*, 40(1), 176-197. <https://doi.org/10.1093/applin/amx020>
- Papi, Z. (2015). Evaluation of Prospect 1 in terms of task types. *International Journal of Humanities and Social Sciences*, 5(2), 238-244.
- Peng, Y., & Pyper, J. S. (2021). Finding success with pedagogical innovation: A case from CSL teachers' experiences with TBLT. *Language Teaching Research*, 25(4), 633-655.
- Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language research and instruction. In S. M. Gass & G. Crookes (Eds.), *Task-based learning in a second language* (pp. 9-34). Multilingual Matters.
- Prabhu, N. S. (1987). *Second language pedagogy*. Oxford University Press.
- Rahimpour, H., & Kiani, G. R. (2015). The difference between rural and urban EFL learners' achievement through different types of task. *Journal of Applied Linguistics and Language Research*, 2(4), 168-178.
- Rebahan, Z. (2019). *Yabancı dil olarak Türkçe öğretiminde kullanılan ders kitaplarındaki konuşma etkinliklerinin göreve dayalı dil öğretimine göre incelenmesi* [Unpublished master's thesis]. Muğla Sıtkı Koçman University.
- Révész, A. (n.d.). The role of the tasks in the syllabus [MOOC Lecture]. In A. Paran, A. Révész, & M. Domingo, *Teaching EFL/ESL reading: A task based approach*. Coursera. <https://www.coursera.org/learn/esl-reading/home/welcome>

- Révész, A. J. (2007). *Focus on form in task-based language teaching: Recasts, task complexity, and L2 learning* [Unpublished doctoral dissertation]. Columbia University.
- Révész, A. (2009). Task complexity, focus on form, and second language development. *Studies in Second Language Acquisition*, 31(3), 437-470.
<https://doi.org/10.1017/S0272263109090366>
- Révész, A. (2011). Task complexity, focus on L2 constructions, and individual differences: A classroom-based study. *The Modern Language Journal*, 95(1), 162–181. <https://doi.org/10.1111/j.1540-4781.2011.01241.x>
- Révész, A. (2019). Replication in task-based language teaching research: Kim (2012) and Shintani (2012). *Language Teaching*, 52(3), 374-384.
- Révész, A., Sachs, R., & Mackey, A. (2011). Task complexity, uptake of recasts, and L2 development. In P. Robinson (Ed.), *Second language task complexity* (pp. 203–235). John Benjamins.
- Riccardi, D. (2014). *Task complexity, task type, and learner-learner interaction: a replication study with adult ESL learners* [Unpublished doctoral dissertation]. University of Toronto.
- Richards, J. C. (2005). *Communicative language teaching today*. SEAMEO Regional Language Centre.
- Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning*, 45(1), 99-140.
- Robinson, P. (2001a). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287-318). Cambridge Applied Linguistics.

- Robinson, P. (2001b). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22(1), 27–57.
<https://doi.org/10.1093/applin/22.1.27>
- Robinson, P. (2003). The cognitive hypothesis, task design and adult task-based language learning. *Second Language Studies*, 21(2), 45–107.
- Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics in Language Teaching (IRAL)*, 43(1), 45–107.
<https://doi.org/10.1515/iral.2005.43.1.1>
- Robinson, P. (2007). Criteria for grading and sequencing pedagogic tasks. In M. D. P. Garcia-Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 7–27). Multilingual Matters.
- Robinson, P. (2008). *Task complexity, intentional reasoning demands, L2 speech production, learning and syllabus design*. LAUD Papers.
- Robinson, P. (2010). Situating and distributing cognition across task demands: The SSARC model of pedagogic task sequencing. In Martin Putz & Laura Sicola (Eds.), *Cognitive processing in second language acquisition: Inside the learner's mind*, (pp. 239-264). John Benjamins.
- Robinson, P. (2011). Second language task complexity, the Cognition Hypothesis, language learning, and performance. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 3-37). John Benjamins.
- Robinson, P. (2015). The Cognition Hypothesis, second language task demands and the SSARC model of pedagogic task sequencing. In M. Bygate (Ed.), *Domains and*

directions in the development of TBLT: Plenaries from a decade of the international conference (pp 87-133). John Benjamins.

- Robinson, P. (2022). The Cognition Hypothesis, the Triadic Componential Framework and the SSARC model: An instructional design theory of pedagogic task sequencing. In M. Ahmadian & M. H. Long (Eds.), *The Cambridge Handbook of Task-Based Language Teaching* (pp. 205-225). Cambridge University Press.
- Robinson, P., & Gilabert, R. (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. *International Review of Applied Linguistics in Language Teaching (IRAL)*, 45(3), 161-176.
<https://doi.org/10.1515/iral.2007.007>
- Robinson, P., Ting, S. C., & Urwin, J. (1996). Three dimensions of second language task complexity. *The University of Queensland Working Papers in Language and Linguistics*, 1(1), 16-32.
- Rodríguez Sánchez, J. (2018). *Language teaching: Task-supported textbook redesign* [Unpublished master's thesis]. Universitat de Barcelona.
- Şahin, M. (2019). *Görev Temelli Dil Öğretim Yöntemi'nin Türkçeyi yabancı dil olarak öğrenenlerin konuşma becerilerinin geliştirilmesinde kullanımı* [Unpublished master's thesis]. Nevşehir Hacı Bektaş Veli University.
- Samuda, V., & Bygate, M. (2008). *Tasks in second language learning*. Palgrave Macmillan.
- Sasayama, S. (2015). *Validating the assumed relationship between task design, cognitive complexity, and second language task performance* [Unpublished master's thesis]. Georgetown University.

- Sasayama, S. (2016). Is a “complex” task really complex? Validating the assumption of cognitive task complexity. *The Modern Language Journal*, 100(1), 231–254.
<https://doi.org/10.1111/modl.12313>
- Scriven, M. (2005). *The logic and methodology of checklists*.
<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.588.7093&rep=rep1&type=pdf>
- Selim, İ. (2015). *Görev odaklı dil öğretimi ile Arapça dil bilgisi öğretimi* [Unpublished master’s thesis]. Gazi University.
- Shadabi, S., Golshan, M., & Sayadian, S. (2017). The effect of one-way and two-way tasks on Iranian EFL learners’ lexical learning. *Journal of Language Teaching and Research*, 8(5), 953-959. <http://dx.doi.org/10.17507/jltr.0805.16>
- Shintani, N. (2012). Input-based tasks and the acquisition of vocabulary and grammar: A process-product study. *Language Teaching Research*, 16(2), 253-279.
- Shintani, N. (2016). *Input-based tasks in foreign language instruction for young learners*. John Benjamins.
- Şimşek, B., & Bakir, S. (2019). The use of task-based language teaching method to teach terms and phrases for those learning Turkish as a second language and sample activities. *Journal of Language and Linguistic Studies* 15(2), 719-738.
- Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied linguistics*, 17(1), 38-62. <https://doi.org/10.1093/applin/17.1.38>
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510-532.
<https://doi.org/10.1093/applin/amp047>

- Skehan, P. (2015). Limited Attention Capacity and cognition: Two hypotheses regarding second language performance on tasks. In M. Bygate (Ed.), *Domains and directions in the development of TBLT: Plenaries from a decade of the international conference* (pp 123-156). John Benjamins.
- Skehan, P. (2016). Tasks versus conditions: Two perspectives on task research and their implications for pedagogy. *Annual Review of Applied Linguistics*, 36, 34-49.
- Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language learning*, 49(1), 93-120.
<https://doi.org/10.1111/1467-9922.00071>
- Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183-205). Cambridge University Press.
- Skehan, P., & Foster, P. (2005). Strategic and on-line planning: The influence of surprise information and task time on second language performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 193-216). John Benjamins. <https://doi.org/10.1075/llt.11>
- Skehan, P., & Foster, P. (2012). Complexity, accuracy, fluency and lexis in task-based performance. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 199–220). John Benjamins.
- Smith, D. B. (2001). *Taking students to task: Task-based computer-mediated communication and negotiated interaction in the ESL classroom* [Unpublished doctoral dissertation]. The University of Arizona.
- Solon, M., Long, A. Y., & Gurzynski-Weiss, L. (2017). Task complexity, language-related episodes, and production of L2 Spanish vowels. *Studies in Second*

Language Acquisition, 39(2), 347-380.

<https://doi.org/10.1017/S0272263116000425>

Su-Bergil, A. & Öcel, M. (2022). İngilizcenin yabancı dil olarak öğretildiği bir ortamda öğretmenlerin görev temelli dil öğretimine karşı tutumları. *Sakarya Üniversitesi Eğitim Fakültesi Dergisi*, 22(1), 68-82.

Sülükçü, D. (2010). *A study of cognitive task complexity and written output at upper-intermediate learners of English* [Unpublished master's thesis]. Hacettepe University.

Tabari, A. (2018). *The effects of planning time conditions on complexity, accuracy, and fluency in second language written narratives* [Unpublished doctoral dissertation]. Oklahoma State University.

Tarone, E. (2013). Interlanguage. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (2794–2801). Wiley-Blackwell.

Terekhova, A. (2017). *A comparative analysis of two Spanish and Russian high school English textbooks* [Unpublished master's thesis]. Universidad de Castilla-La Mancha.

Un-udom, S., & Patanasorn, C. (2020). Effects of task sequencing in the Cognition Hypothesis on EFL learner perception of task difficulty. *Journal of Studies in the English Language*, 15(2), 1-26.

Uzun, L., & Ertok, Ş. A. (2020). Student opinions on task-based approach as formative evaluation versus exam-based approach as summative evaluation in education. *Sakarya University Journal of Education*, 10(2), 226-250.

<https://doi.org/10.19126/suje.598048>

Vasylets, O. (2017). *Task-modality effects: A study of task complexity effects in speech and writing* [Unpublished doctoral dissertation]. Universitat de Barcelona.

- Vasylets, O., Gilabert, R., & Manchon, R. M. (2017). The effects of mode and task complexity on second language production. *Language Learning*, 67(2), 394-430.
<https://doi.org/10.1111/lang.12228>
- Vo, Y. (2018). *Effects of integrated and independent speaking tasks on learners' interactional performance* [Unpublished doctoral dissertation]. Minnesota State University.
- Wang, Z., & Skehan, P. (2014). Structure, lexis, and time perspective. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 155-186).
<https://doi.org/10.1075/tblt.5.06wan>
- Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford University Press.
- Willis, J. (1996). *A framework for task-based learning*. Longman.
- Yang, W. (2014). *Mapping the relationships among the cognitive complexity of independent writing tasks, L2 writing quality, and complexity, accuracy and fluency of L2 writing* [Unpublished doctoral dissertation]. Georgia State University.
- Yıldız, M. (2012). *Teaching grammar through task-based language teaching to young EFL learners* [Unpublished master's thesis]. Ondokuz Mayıs University.
- Yıldız, M. (2017). *Effects of task complexity and rhetorical mode on writing performance of EFL learners* [Unpublished doctoral dissertation]. Atatürk University.
- Yoon, H. J. (2017). *Investigating the interactions among genre, task complexity, and proficiency in L2 writing: A comprehensive text analysis and study of learner perceptions* [Unpublished doctoral dissertation]. Michigan State University.

- York, J. (2019). *Language learning in complex virtual worlds: Effects of modality and task complexity on oral performance between virtual world and face-to-face tasks* [Unpublished doctoral dissertation]. University of Leicester.
- Yuan, F. (2001). *The effects of planning on language production in task-based language teaching* [Unpublished doctoral dissertation]. Temple University.
- Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, 24(1), 1-27. <https://doi.org/10.1093/applin/24.1.1>
- Zhu, Y. (2020). Implementing tasks in young learners' language classrooms: A collaborative teacher education initiative through task evaluation. *Language Teaching Research*, 00(0), 1-22. <https://doi.org/10.1177/1362168819894706>
- Zohrabi, M., & Hassanpour, S. (2020). The effect of open vs. closed tasks on improving Iranian EFL learners' oral performance. *Journal of Teaching Language Skills (JTLS)* 38(4), 205-238.

Appendix

Development and Analysis of a Sample Task

This section provides examples of a non-task activity and a sample task. The non-task activity, taken from the *Silver Lining* for grade 11, meets some of the task criteria defined by Ellis (2009) and Ellis and Shintani (2014) and explicated by Jaruteerapan (2020). The sample task was developed by redesigning this non-task activity to satisfy the task criteria. First, the task characteristics of the non-task activity and the sample task will be examined. Then, the sample task will be further analysed in terms of the task types defined by Ellis (2003, 2012, 2020). Finally, the task complexity dimensions of the sample task will be discussed with specific regard to the SSARC model by Robinson (2001a, 2001b, 2003, 2007, 2008, 2011, 2015).

Task Characteristics

Non-task Activity

Theme 8 Activity 7 in Silver Lining for Grade 11

7 Work in pairs. Choose one of the tournaments and talk to your partner using the prompts.

Formula 1 Race



- single-seat auto racing
- 20 racers
- lasts for three days
- open to men and women
- usually starts with practices on Friday and finishes with a race on Sunday

Istanbul Eurasia Marathon



- lasts for one day
- from 9 a.m. to 3 p.m.
- starts on the Asian part of the city, crosses the Bosphorus Bridge and finishes on the European part
- online application
- 3000 participant limit

Formula 1 is a single-seat auto race. 20 drivers race for the cup. It

Note. From *Silver Lining 11 (Student's Book)* (p. 102), by E. Akdağ, F. Baydar

Ertopçu, K. Uyanık Bektaş, S. Umur Özadalı, & T. Kaya, 2019, Milli Eğitim

Bakanlığı. Copyright 2019 by Milli Eğitim Bakanlığı.

The activity above satisfies criterion 1 (meaning) as learners “mainly use language for a communicative purpose” and are “primarily concerned with producing and/or comprehending the messages (semantic and pragmatic meaning), not with focusing one linguistic form” (Jaruteerapan, 2020, p. 71). As for criterion 3, they mainly depend on their own linguistic resources to implement the activity. Even if the information they are to share is already given and there is some linguistic starting point provided to discuss the tournaments, there is no teaching of the language to be used, and learners are still allowed to use the language they already know. However, the activity fails to meet criteria 2 (gap) and 4 (outcome). First, as learners can see the information for each tournament, there is no genuine gap (criterion 2) that necessitates the transfer of a message and gets closed when learners participate in communication (Jaruteerapan, 2020). Second, there is not a clearly defined outcome (criterion 4) such as a “communicative outcome”, “abstract idea”, or “physical reactions” (Jaruteerapan, 2020, p. 72) that would signal the completion of the activity. Learners are merely instructed to talk to one another using the prompts without a communicative goal.

The activity above can be called *task-like* as it meets some of the criteria of taskness proposed by Ellis (2009) and Ellis and Shintani (2014) and explained by Jaruteerapan (2020). Below is a sample task proposed by the researcher by redesigning this task-like activity to satisfy the task criteria in question.

Sample Task

You and your partner are two sports fans planning to attend a sports event for a weekend. You both have searched for the upcoming tournaments in İstanbul in the fall and took some notes on your phones to suggest an event to your partners.

Student A: Turn to page 135

Student B: Turn to page 140

Page 135

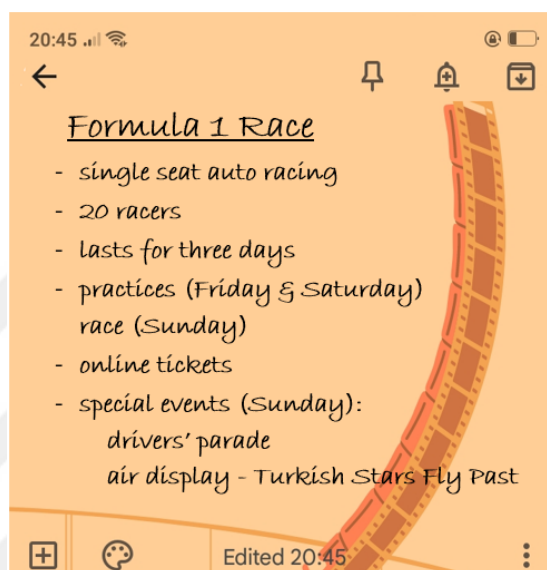
Student A:

Read your notes about the *Formula 1 Race* and prepare for two minutes to share them with your partner.

Give information about:

- what the event is for
- where/when it takes place
- how long it lasts
- how to attend
- special event(s)

Compare your partner's suggestion with yours and discuss which one you would like to attend by explaining your reasons. Then, decide on the event to attend together with your partner.



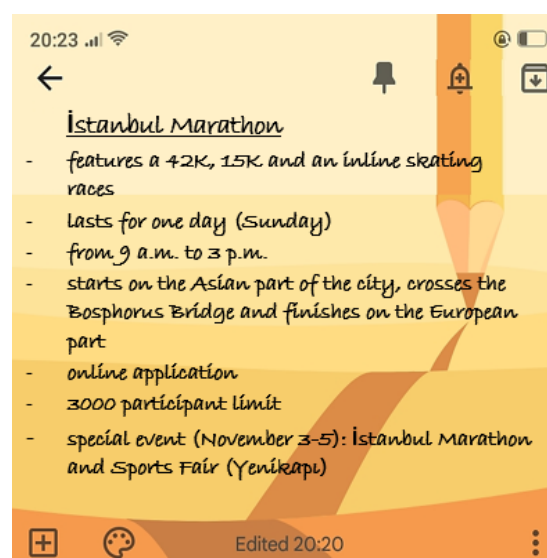
Page 140

Student B:

Read your notes about the *İstanbul Marathon* and prepare for two minutes to share them with your partner. Give information about:

- what the event is for
- where/when it takes place
- how long it lasts
- how to attend
- special event(s)

Compare your partner's suggestion with yours and discuss which one you would like to attend by explaining your reasons. Then, decide on the event to attend together with your partner.



The sample task was generated based on the non-task activity provided in the EFL textbook for grade 11, *Silver Lining* (Akdağ et al., 2019, p. 102). The researcher redesigned the task content based on the tournaments to take place in İstanbul and revised the task instructions accordingly to increase situational authenticity. With this, the researcher also aimed to strengthen the activity regarding both *gap* (criterion 2) and *outcome* (criterion 4) aspects.

The sample task introduces a communicative *gap* (criterion 2) in two ways. First, it creates a real gap which motivates learners to communicate to determine which event to attend. Second, the task content is provided on two different pages to be read by different partners so that there would be a genuine need to transfer the message from one person to another and fill the gap.

As for criterion 4 (outcome), the sample task provides learners with a communicative goal that would go beyond the mere display of linguistic knowledge. They need to work on the task content to collaboratively decide what tournament to attend together, the completion of which also yields a communicative outcome and signals the end of the task.

Please note that the instruction “talk to your partner using the task prompt” (Akdağ et al., 2019, p. 102) in the non-task activity was replaced with other verbs such as *share*, *compare*, *discuss*, *explain* and *decide on*, which help structure the task and easily walk learners through the task steps. The last verb *decide on* in the instructions ensures that learners work towards a communicative outcome, “abstract idea”, as named by Jaruteerapan (2020, p. 72).

Task Types

The sample task will be examined based on the five dual categories proposed by Ellis (2003, 2012, 2020) for task typology.

First, the sample task does not orient learners to the use of certain linguistic features but allows them to use language communicatively. Therefore, it can be considered an example of an unfocused (UN) task.

The sample task has situational authenticity as it imitates real-life context, in which learners may search for sports events online, take some notes on their mobile phones and share them with their friends to determine which event to attend together outside school. Thus, the language produced during the task implementation is more similar to the one occurring in non-pedagogic conversations as in real-world (RW) tasks.

As for the receptive or productive skills which the task engages learners with, the task can be called an integrative (INT) task as it requires the use of skills from both input-based (IB) and output-based (OB) tasks.

When it comes to the presence or absence of task contents during task implementation, the contents of the task are present in front of the learners, and learners can revisit them whenever they want, as in here-and-now (HN) tasks.

Finally, considering the scope of the task outcome, the sample task is an open (OP) task as it does not have one correct outcome but allows learners to make their own choices.

Task Complexity

The sample task is predicted to be complex along there-and-then (TT) and single step (SS) dimensions based on the task complexity framework by Robinson (2001a, 2001b, 2003, 2007, 2008, 2011, 2015). The task necessitates displaced references to space and time (+TT), which increases cognitive/conceptual complexity. Besides, more than one step is involved in the tasks to accomplish the

task outcomes (-SS), which increases the performative/procedural demands during task implementation.

Considering that the TT variable is a resource-directing dimension, and the SS is a resource-dispersing dimension, it can be claimed that the task is complex along both resource-directing and resource-dispersing dimensions (+RDir +RDis), which corresponds to the third, and the most complex, step of the SSARC model.

