

Informal second language learning:
The role of engagement, proficiency,
attitudes, and motivation



Henriette L. Arndt

Department of Education

St. Hilda's College

University of Oxford

Thesis submitted for the degree of

Doctor of Philosophy

Trinity Term 2019

(90,887 words, excluding references and appendices)

Thesis Abstract

The majority of research in second language acquisition has traditionally focused on classroom-based, instructed learning. However, with the spread of new technologies and the concomitant globalisation of popular culture, informal second language practices (ISLPs; e.g. watching films and television, using social media, or playing video games in another language) are increasingly becoming part of the daily lives of many language learners, particularly those who are learning English. Previous studies of *Informal Second Language Learning* (ISLL) have provided evidence for a positive relationship between such informal language contact and L2 proficiency. However, considerable individual and group differences have also been recorded, which seem to indicate that not all learners participate in and/or can benefit equally from these activities.

The present study employed a framework of student engagement (comprising behavioural, cognitive, affective, and linguistic aspects) in an attempt to provide deeper insight into the nature of German secondary school students' informal second language practices. It also explored the relationships between informal language contact and proficiency, motivation, and attitudes towards language learning, as well as their directionality; that is, the extent to which ISLPs promoted language development and changes in motivation, and whether more proficient and motivated students were more likely to engage in informal activities.

The project involved mixed research methods with data collected in three phases: In Phase I, the researcher conducted focus group interviews with 47 students, which were subsequently transcribed and analysed via inductive and deductive coding. The qualitative findings then informed the development of an event-contingent diary for recording informal second language practices, including items to measure the different aspects of engagement, and a questionnaire for studying motivation and attitudes specifically in the context of ISLL. At this point, a C-test was also developed as a measure of general language proficiency in this particular learner population. After extensive piloting (Phase II) with a sample of 105 students, the new instruments were used to collect data from 354 further participants at two time points in Phase III. The nature of the students' engagement in informal second language practices and its relationships with motivation, attitudes, and proficiency over time were explored through statistical modelling (Multi-level Structural Equation Modelling and Latent Profile Analysis) and triangulated with the qualitative findings from Phase I.

This study provided novel insights into the nature of learners' interactions with informal second language practices. Whereas previous research has primarily focused on the time learners spend on informal practices and the types of activities they choose (behavioural engagement), the current study also considered the role of emotions and attention to meaning and form (affective, cognitive, and linguistic

engagement). Furthermore, the study yielded statistical evidence of the significant positive link between engagement in ISLPs and proficiency, which has been widely reported in the ISLL literature. Meanwhile, the findings, have also called into question the directionality of this relationship: The participants' prior proficiency explained a significant proportion of the variance in the overall engagement patterns, as well as some of the individual aspects of engagement (behavioural, cognitive, and linguistic engagement). However, none of the engagement measures significantly predicted changes in general language proficiency over the course of the study.

By contrast, there was evidence of a reciprocal relationship between motivation and engagement in informal second language practices. That is, engagement was significantly predicted by learners' attitudes towards informal and formal language learning, their desire to become successful second language users (ideal L2 selves), and their reasons for wanting to learn English (language learning orientations). In turn, some aspects of engagement in ISLPs also predicted changes in students' beliefs about formal language learning, their perceptions of themselves as language users, and the extent to which they wanted to learn the language (as opposed to being obligated to do so).

Altogether, this research has made strong theoretical, empirical, and methodological contributions to the emerging field of Informal Second Language Learning, as well as the wider field of Applied Linguistics. The findings are also of practical interest to, for example, teachers and policy makers seeking to gain a better understanding of the factors that determine participation and success in language learning beyond the classroom.

Acknowledgements

First and foremost, I would like to give thanks to all of the pupils who volunteered to take part in this study and to their teachers, who dedicated great amounts of time and effort towards making this research possible. I also gratefully acknowledge the funding I received from the UK Economic and Social Research Council, the Foundation of German Business, as well as the Department of Education and St. Hilda's College at the University of Oxford.

I owe a great deal of gratitude to my supervisors, Dr Robert Woore and Dr Jessica Briggs Baffoe-Djan, for being outstanding teachers and mentors. They have provided me with countless insightful comments and suggestions, which have greatly helped in shaping the direction of my work, and I am sincerely grateful for their unwavering support, compassion, and encouragement.

I want to thank Prof Victoria Murphy, Prof Lars-Erik Malmberg, and Dr Heath Rose for the feedback I have received from them about my research. I am also grateful to Prof Geoffrey Sockett, Dr Meryl Kusyk, Dr Denyze Toffoli, Prof Liss-Kerstin Sylvén, and other members of the *Informal Second Language Learning Network*, who have inspired me through their work and the discussions we have had. A special thanks also to Dr Janina Iwaniec and Arnold Kochari for taking the time to share their specialist knowledge of Applied Linguistics theories and research methods with me.

I would also like to express my greatest appreciation to my friends for all they have done to help me. Ashmita Randhawa, Anna-Maria Ramezanzadeh, Kari Sahan, Jaimie Miller-Friedmann, Vânia Pinto, and Emma Abotsi: Thank you for all of your emotional and academic support and many much-needed coffee breaks. You have been absolutely invaluable in keeping me motivated and grounded throughout my DPhil.

Finally, I am deeply grateful to my parents, Elfriede and Reinhard Arndt, and my brother, Nicolas Arndt, for always believing in me. Without their support I would not be where I am today.

Table of Contents

THESIS ABSTRACT	I
ACKNOWLEDGEMENTS	III
TABLE OF CONTENTS	IV
LIST OF ABBREVIATIONS	VIII
1. INTRODUCTION	1
2. LITERATURE REVIEW	7
2.1 INFORMAL SECOND LANGUAGE LEARNING	7
2.1.1 <i>Describing Informal Second Language Practices</i>	11
2.1.2 <i>Identifying Outcomes of Informal Second Language Learning</i>	17
2.2 ENGAGEMENT	31
2.2.1 <i>Theoretical Background</i>	31
2.2.2 <i>Engagement in Informal Second Language Learning</i>	38
2.3 MOTIVATION AND ATTITUDES	44
2.3.1 <i>Theoretical Background</i>	45
2.3.2 <i>Motivation and Attitudes in Informal Second Language Learning</i>	57
3. METHODOLOGY: INTRODUCTION TO THE STUDY	65
3.1 THEORETICAL MODEL AND RESEARCH QUESTIONS	65
3.2 RESEARCH DESIGN	66
3.3 TARGET POPULATION AND RECRUITMENT	71
3.4 ETHICAL CONSIDERATIONS	74
4. METHODS PHASE I: QUALITATIVE DATA COLLECTION	77
4.1 PARTICIPANTS	77
4.2 PROCEDURE	78
4.3 ANALYSIS	82
5. FINDINGS PHASE I: QUALITATIVE FINDINGS	85
5.1 ENGAGEMENT: HOW AND WHY DO THESE STUDENTS ENGAGE IN INFORMAL SECOND LANGUAGE PRACTICES?	85
5.1.1 <i>Affective Engagement</i>	96
5.1.2 <i>Cognitive Engagement</i>	98
5.1.3 <i>Linguistic Engagement</i>	100
5.2 MOTIVATION FOR LANGUAGE LEARNING	103
5.3 ATTITUDES TOWARDS LANGUAGE LEARNING	108
5.3.1 <i>Perceived Learning Gains</i>	109
5.3.2 <i>Types of Knowledge Gained</i>	113
5.3.3 <i>(In)Formal Learning and Affective States</i>	117
5.4 KEY FINDINGS	118
6. METHODS PHASE II: INSTRUMENT DEVELOPMENT AND PILOTING	122
6.1 INITIAL INSTRUMENT DEVELOPMENT	122
6.1.1 <i>C-tests</i>	122
6.1.2 <i>Motivation and Attitudes Questionnaire</i>	126
6.1.3 <i>Informal Language Learning Diary</i>	129
6.2 PILOTING: PROCEDURE	135
6.2.1 <i>A Note on Participant Numbers and Test Time During the C-tests</i>	136
6.2.2 <i>C-Test Scoring</i>	138
6.2.3 <i>Data Analysis</i>	139
7. FINDINGS PHASE II: EVALUATING AND FINALISING THE INSTRUMENTS	142
7.1 C-TESTS	142
7.1.1 <i>Descriptive Statistics and Equality of Means</i>	142

7.1.2	<i>Normality and Outliers</i>	143
7.1.3	<i>Validity: Exploratory Factor Analysis</i>	145
7.1.4	<i>Reliability: Cronbach's Alpha</i>	147
7.1.5	<i>Item Exclusions</i>	148
7.1.6	<i>Adding More Difficult Items</i>	149
7.2	MOTIVATION AND ATTITUDES QUESTIONNAIRE	150
7.3	INFORMAL LANGUAGE LEARNING DIARY	156
7.3.1	<i>Participant Compliance</i>	160
8.	METHODS PHASE III: LARGE-SCALE QUANTITATIVE DATA COLLECTION	163
8.1	PARTICIPANTS	163
8.2	PROCEDURE	164
8.3	DATA ANALYSIS	165
9.	FINDINGS PHASE III: QUANTITATIVE ANALYSES	167
9.1	PRELIMINARY ANALYSES	167
9.1.1	<i>Completeness of Data</i>	167
9.1.2	<i>Normality and Outliers</i>	168
9.1.3	<i>Validity: Confirmatory Factor Analysis and Measurement Invariance Modelling</i>	169
9.1.4	<i>Reliability: Cronbach's Alpha</i>	196
9.2	RESEARCH QUESTION 1: HOW DO GERMAN LEARNERS OF ENGLISH AS A SECOND LANGUAGE ATTENDING GYMNASIEN ENGAGE IN ENGLISH INFORMAL SECOND LANGUAGE PRACTICES?	198
9.2.1	<i>Quantity and Variety of Informal Second Language Practices</i>	198
9.2.2	<i>Popularity of Different Informal Second Language Activities</i>	204
9.2.3	<i>Levels of Engagement in Different Informal Second Language Activities</i>	210
9.2.4	<i>Relationships Between the Different Aspects of Engagement</i>	214
9.2.5	<i>Are There Distinct Overall Patterns of Engagement in ISLPs Among the Participants?</i>	216
9.2.6	<i>Key Findings</i>	222
9.3	RESEARCH QUESTION 2: WHAT IS THE DIRECT RELATIONSHIP BETWEEN GENERAL LANGUAGE PROFICIENCY AND ENGAGEMENT IN INFORMAL SECOND LANGUAGE PRACTICES?	224
9.3.1	<i>Proficiency and Types of Informal Second Language Activities</i>	225
9.3.2	<i>Structural Equation Models and Path Modelling</i>	227
9.3.3	<i>Proficiency and the Different Dimensions of Engagement</i>	229
9.3.4	<i>Proficiency and Engagement Profiles</i>	233
9.3.5	<i>Key Findings</i>	235
9.4	RESEARCH QUESTION 3: WHAT ARE THE DIRECT RELATIONSHIPS BETWEEN LANGUAGE LEARNING MOTIVATION AND ATTITUDES TOWARDS LANGUAGE LEARNING AND ENGAGEMENT IN INFORMAL SECOND LANGUAGE PRACTICES?	237
9.4.1	<i>Motivation, Attitudes and Dimensions of Engagement</i>	238
9.4.2	<i>Motivation, Attitudes, and Engagement Profiles</i>	244
9.4.3	<i>Key Findings</i>	246
9.5	RESEARCH QUESTION 4: HOW DO GENERAL LANGUAGE PROFICIENCY, LANGUAGE LEARNING MOTIVATION, AND ATTITUDES INTERACT IN THEIR RELATIONSHIPS WITH ENGAGEMENT IN INFORMAL SECOND LANGUAGE PRACTICES?	249
9.5.1	<i>Behavioural, Affective, Cognitive, and Linguistic Engagement</i>	250
9.5.2	<i>Engagement Profiles</i>	258
9.5.3	<i>Key Findings</i>	260
10.	DISCUSSION	263
10.1	ENGAGEMENT IN INFORMAL SECOND LANGUAGE PRACTICES	263
10.1.1	<i>Behavioural Engagement</i>	264
10.1.2	<i>Affective, Cognitive, and Linguistic Engagement</i>	269
10.1.3	<i>Relationships between Engagement Dimensions</i>	274
10.1.4	<i>Engagement Profiles</i>	279
10.2	WHAT ARE THE RELATIONSHIPS BETWEEN GENERAL LANGUAGE PROFICIENCY AND ENGAGEMENT IN INFORMAL SECOND LANGUAGE PRACTICES?	283
10.2.1	<i>Do Informal Second Language Practices Promote Proficiency?</i>	283
10.2.2	<i>Relationships Between Proficiency and Various Aspects of Engagement</i>	293

10.3	WHAT ARE THE RELATIONSHIPS BETWEEN MOTIVATION AND ATTITUDES TOWARDS LANGUAGE LEARNING AND ENGAGEMENT IN INFORMAL SECOND LANGUAGE PRACTICES?	304
10.3.1	<i>Language Learning Orientations</i>	305
10.3.2	<i>Second Language Motivational Selves</i>	313
10.3.3	<i>Attitudes towards Informal and Formal Learning</i>	317
11.	CONCLUSION	322
11.1	KEY FINDINGS	322
11.2	STRENGTHS OF THE STUDY	328
11.3	LIMITATIONS	330
11.4	DIRECTIONS FOR FUTURE RESEARCH	335
11.5	IMPLICATIONS FOR TEACHING AND LEARNING	338
	REFERENCES	345
APPENDIX A.	METHODS PHASE I	357
A.1	FOCUS GROUP INTERVIEW GUIDE	357
A.1.1	<i>English</i>	357
A.1.2	<i>German</i>	360
A.2	PARTICIPANT QUESTIONNAIRE	363
A.2.1	<i>English</i>	363
A.2.2	<i>German</i>	367
APPENDIX B.	METHODS PHASE II AND III	371
B.1	C-TESTS	371
B.1.1	<i>Version 1</i>	372
B.1.2	<i>Version 2</i>	374
B.2	MOTIVATION AND ATTITUDES QUESTIONNAIRE	377
B.2.1	<i>Overview of Motivation and Attitudes Constructs and Questionnaire Items</i>	377
B.2.2	<i>English Questionnaire</i>	380
B.2.3	<i>German Questionnaire</i>	391
B.3	INFORMAL SECOND LANGUAGE LEARNING DIARY	402
B.3.1	<i>Overview of Engagement Constructs and Survey Items</i>	402
B.3.2	<i>English Instructions</i>	403
B.3.3	<i>German Instructions</i>	406
B.3.4	<i>English List of Questions</i>	409
B.3.5	<i>German List of Questions</i>	414
B.3.6	<i>Screenshots of the Online Survey</i>	419
APPENDIX C.	FINDINGS PHASE I	421
C.1	QUALITATIVE CODING FRAMEWORK	421
APPENDIX D.	FINDINGS PHASE II	423
D.1	C-TESTS PRELIMINARY ANALYSES: NORMALITY AND OUTLIERS	423
D.2	C-TESTS: LEXICAL DIVERSITY AND READABILITY STATISTICS	424
D.3	EXPLORATORY FACTOR ANALYSES	426
D.3.1	<i>C-Tests</i>	426
D.3.2	<i>Language Learning Orientations</i>	427
D.3.3	<i>Second Language Motivational Selves</i>	430
D.3.4	<i>Attitudes Towards Informal and Formal Language Learning</i>	431
D.3.5	<i>Engagement</i>	432
APPENDIX E.	FINDINGS PHASE III	433
E.1	PRELIMINARY ANALYSES: NORMALITY AND OUTLIERS	433
E.2	EFFECTS OF CLASSROOM CLUSTERING	440
E.3	CONFIRMATORY FACTOR ANALYSIS: C-TESTS	440

E.4	MEASUREMENT INVARIANCE MODELLING: C-TESTS	442
E.5	CONFIRMATORY FACTOR ANALYSIS: ENGAGEMENT	447
APPENDIX F. FINDINGS PHASE III, RQ 1: HOW DO THESE LEARNERS ENGAGE IN ENGLISH INFORMAL SECOND LANGUAGE PRACTICES?		449
F.1	DERIVING MEASURES OF BEHAVIOURAL ENGAGEMENT	449
F.2	POPULARITY OF DIFFERENT ACTIVITIES	450
F.3	FINDINGS REGARDING SPECIFIC INFORMAL SECOND LANGUAGE PRACTICES	454
F.3.1	<i>Film and Television</i>	454
F.3.2	<i>Social Media</i>	455
F.3.3	<i>Writing</i>	458
F.3.4	<i>Speaking</i>	458
F.3.5	<i>Music</i>	459
F.3.6	<i>Video Games</i>	460
F.3.7	<i>Reading</i>	461
F.3.8	<i>Listening</i>	462
F.4	GENDER AND TYPES OF SECOND LANGUAGE ACTIVITIES	463
F.5	ENGAGEMENT PROFILES AND TYPES OF SECOND LANGUAGE ACTIVITIES	464
APPENDIX G. FINDINGS PHASE III, RQ 2: WHAT ARE THE DIRECT RELATIONSHIPS BETWEEN GENERAL LANGUAGE PROFICIENCY AND ENGAGEMENT IN INFORMAL SECOND LANGUAGE PRACTICES?		467
G.1	PROFICIENCY AND TYPES OF SECOND LANGUAGE ACTIVITIES	467
G.2	PROFICIENCY AND DIMENSIONS OF ENGAGEMENT	469
APPENDIX H. FINDINGS PHASE III RQ 3: WHAT ARE THE DIRECT RELATIONSHIPS BETWEEN LANGUAGE LEARNING MOTIVATION AND ATTITUDES TOWARDS LANGUAGE LEARNING AND ENGAGEMENT IN INFORMAL SECOND LANGUAGE PRACTICES?		472
H.1	MOTIVATION, ATTITUDES AND DIMENSIONS OF ENGAGEMENT	472
H.2	MOTIVATION, ATTITUDES AND ENGAGEMENT PROFILES	474
APPENDIX I. FINDINGS PHASE III RQ 4: HOW DO GENERAL LANGUAGE PROFICIENCY, LANGUAGE LEARNING MOTIVATION, AND ATTITUDES INTERACT IN THEIR RELATIONSHIPS WITH ENGAGEMENT IN INFORMAL SECOND LANGUAGE PRACTICES?		475
I.1	BEHAVIOURAL, AFFECTIVE, COGNITIVE, AND LINGUISTIC ENGAGEMENT	475
I.2	ENGAGEMENT PROFILES	478

List of Abbreviations

ANOVA	Analysis of variance
CFA	Confirmatory factor analysis
CFI	Comparative fit index
CTLs	Classroom-trained learners (Cole, 2015)
EFA	Exploratory factor analysis
EFL	English as a foreign language
ESL	English as a second language
FASILs	Fully autonomous self-instructed learners (Cole, 2015)
FIML	Full information maximum likelihood modelling
FS	Factor scores
ISLL	Informal second language learning
ISLPs	Informal second language practices
KMO	Kaiser-Meyer-Olkin measure
L1	First language
L2	Second language
L2MSS	Second language motivational self-system
LLO	Language learning orientation
LPA	Latent profile analysis
MI	Modification indices
MIM	Measurement invariance modelling
MLR	Robust maximum likelihood
NLS	New literary studies
RMSEA	Root mean square error of approximation
RQ	Research question
SEM	Structural equation model / Structural equation modelling
SE Model	Socio-Educational Model
SLA	Second language acquisition
SRMR	Standardised root mean square residual
T1 / T2	Time 1 / Time 2
TLI	Tucker-Lewis index
V1 / V2	Version 1 / Version 2

1. Introduction

Research in second language acquisition has traditionally focused on classroom-based, instructed learning (Benson, 2011b; White, 1995) with few exceptions, such as research on language learning in study abroad (see for example Kinginger, 2009). However, recently interest has grown in language learning outside of formal educational contexts (e.g. Benson, 2011b; Nunan & Richards, 2015). In part, this can be attributed to the rapid development and spread of new technologies, ever-increasing access to the internet, and the concomitant emergence of a progressively globalised culture over the past decades, which mean that learners now have more opportunities than ever before to independently develop and practice applying their second language (L2) knowledge.

Many materials are available online which have been specifically developed to support independent, *non-formal* language learning, including, for example, podcast- or video-based courses, grammar websites, and language learning apps. Besides these, the internet also affords learners opportunities for more naturalistic, authentic language contact. They may, for example, engage with a range of media in a variety of languages, including (but not limited to) social networking, news and entertainment websites, films and TV series, music, and video games. Such *informal second language practices* (ISLPs) and their impact on language learners' development are the focus of the current thesis.

Previous studies of *Informal Second Language Learning* (ISLL) have provided evidence for a positive relationship between informal practices and second language proficiency (e.g. De Wilde, Brysbaert, & Eyckmans, 2019; Kuppens, 2007, 2010; J. S. Lee & Dressman, 2017; Olsson, 2011; Sundqvist, 2009, 2011; M. H. Verspoor, de Bot,

& Lowie, 2011), as well as some affective variables, including learner motivation, self-efficacy, confidence, and enjoyment (e.g. Cole, 2015; Lai, Zhu, & Gong, 2015; Lamb & Arisandy, 2019; J. S. Lee, 2019; Lyrigkou, 2016). However, considerable individual and group differences have also been recorded, which seem to indicate that not all learners participate in and/or benefit from informal activities equally (e.g. Cole, 2015; Olsson, 2011; Sundqvist, 2009, 2011; Sundqvist & Sylvén, 2014). In addition, most prior studies have employed correlational methods, so that it is unclear to what extent informal activities promote changes in proficiency and motivation, and to what extent more proficient and motivated students are more likely to engage in informal second language practices.

The current study was designed to explore whether the nature of *engagement* in informal second language practices might help to explain some of the aforementioned individual and group differences in proficiency and motivation associated with ISLPs. Engagement is a concept derived from Educational Psychology (e.g. Appleton, Christenson, Kim, & Reschly, 2006; Boekaerts, 2016; Christenson, Reschly, & Wylie, 2012) which describes the ‘qualities of students’ interactions with learning activities’ (Wang, Fredricks, Ye, Hofkens, & Schall Lin, 2016) in terms of their behaviour, thoughts, and feelings (behavioural, cognitive, and affective engagement). Thus, this study first involved quantitatively and qualitatively mapping the engagement in informal second language practices of a specific population of learners of English as a second language, year 10 students at *Gymnasien*¹ in Germany. The main focus of the

¹ A *Gymnasium* is a type of secondary school in Germany which prepares its students for entry into university or other advanced post-secondary education. *Gymnasien* are comparable to grammar schools and sixth form colleges in the UK.

study was on statistically modelling the links between different aspects of engagement and changes in language proficiency, language learning motivation, and attitudes over time.

To make clear how informal second language learning differs from other forms of learning beyond the classroom, the literature review (Chapter 2) opens with a detailed definition of the concept. Following this is a critical summary of the findings of previous studies which (a) describe the nature of informal second language practices and explore their prevalence within different learner populations; and (b) identify their outcomes in terms of second language learning. It is concluded that there is substantial variability in the extent to which learners participate in informal second language activities and in the associated outcomes. To explain this variability, more systematic evidence is needed on *how* and *why* individuals engage informally with their second language. The last part of the literature review thus introduces the concepts of *engagement* and *motivation and attitudes* in language learning as ways to conceptualise, respectively, the nature of informal practices and individuals' reasons for engaging in them. Where these have been addressed in previous ISLL studies, the research findings are subsequently reviewed.

The methodology section (Chapter 3) opens with the proposal of a model which brings together the different elements of the Literature Review (informal second language practices, proficiency, engagement, and motivation and attitudes). At this point, the overall design of the present study is explained, which consisted of three phases involving: (I) exploratory, qualitative data collection through focus group interviews; (II) the development and piloting of a general language proficiency test, a language learning motivation questionnaire, and an informal language learning

diary; and (III) the collection and analysis of quantitative data to test the aforementioned theoretical framework. The methods employed in each phase and the rationale for specific design choices are described in more detail in the respective methods chapters: Chapter 4 for Phase I; Chapter 6 for Phase II, and Chapter 8 for Phase III. The thesis was structured in this way, with separate methods and findings chapters for each phase, to reflect the fact that each new stage of the project built upon the findings from the preceding phases.

Chapter 5 presents the qualitative findings from the semi-structured focus group interviews. The first section focuses on describing the informal second language practices of the sampled students, their reasons for engaging in these activities, as well as reflections on their concurrent thoughts and feelings. The second section addresses the pupils' general motivation for learning English as a second language. The last section discusses the participants' attitudes towards informal and formal learning, their perceptions of the effectiveness of these modes of learning, and the types of knowledge the students believed each mode promoted.

Chapter 7 presents the results from Phase II: the pilot study in which the newly developed instruments were evaluated for their validity and reliability and refined for use in the next phase of this research.

The findings of the large-scale quantitative study which constituted Phase III are presented in Chapter 9. The first section contains results from preliminary analyses, concerning the completeness of the collected data, the normality of the variables and presence of outliers, as well as the reliability and validity of the instruments, taking into account the changes made after piloting. Next, findings are presented which address each of the four research questions at the heart of this project:

1. How do learners of English as a second language, attending year 10 classes at German *Gymnasien*, engage in English informal second language practices?
2. What are the direct relationships between language proficiency and engagement in informal second language practices?
3. What are the direct relationships between language learning motivation and attitudes and engagement in informal second language practices?
4. How do language proficiency, language learning motivation, and attitudes interact in their relationships with engagement in informal second language practices?

The first research question was addressed mainly through descriptive analyses of the data gathered through the informal language learning diary. The primary focus was on the amount of time the participants spent using their second language outside of school, how frequently they did so, the types of practices in which they engaged, and their self-reported levels of affective, cognitive, and linguistic engagement. Alongside these findings, the relevant section also contains the results of a Latent Profile Analysis, which was used to investigate whether groups of learners could be identified who showed distinct patterns of engagement in ISLPs. The rest of the chapter presents the results from a series of Structural Equation Models which were used in the stepwise assessment of the relationships between various aspects of engagement in informal second language practices on the one hand, and L2 proficiency, motivation, and attitudes measured at two time points on the other (research questions 2 through 4).

Chapter 10 contains a discussion of both the qualitative and quantitative findings of this thesis in light of the underpinning theories and earlier research

findings. The thesis concludes (Chapter 11) with a summary of the main results, the study's theoretical and practical strengths and limitations, suggestions for future research directions and a discussion of implications for language learning and teaching more broadly.

2. Literature Review

The main focus of this Chapter is on reviewing the findings of previous research on various aspects of *informal second language learning* (ISLL). It begins with a detailed working definition of ISLL (Section 2.1), before introducing studies which have mapped the popularity of informal second language practices among various populations of second language learners (Section 2.1.1) and which have investigated the relationships between informal activities and various types of language knowledge and skills (Section 2.1.2). It is concluded that, to explain the seemingly complex relationships between informal second language practices and language proficiency, it is necessary to study in more detail *how* and *why* individuals engage in informal second language practices. Thus, the last two sections of the current chapter focus on *engagement* (Section 2.2) and *motivation* and *attitudes* (Section 2.3) as ways to conceptualise, respectively, the qualities of learners' interactions with informal second language practices and their perceptions related to second language learning. Both sections begin with a more general, theoretical review of how these concepts have been treated in the wider literature on Second Language Acquisition (SLA), Psychology, and Education, followed by an overview of relevant findings from previous studies on informal second language learning.

2.1 Informal Second Language Learning

Informal second language learning can be defined by drawing on a framework proposed by Benson (2011a), which distinguishes different forms of language learning beyond the classroom from one another along four 'dimensions':

1. The 'location' or setting in which learning takes place. This serves to distinguish non-prescribed learning, in which students engage independently

(*out-of-class* or *out-of-school learning*), from activities which take place in schools but are not part of regular lessons (*extracurricular* or *after-school* programmes).

2. 'Formality', which is defined by whether or not the learning is aimed at achieving formal qualifications. Thus, this dimension distinguishes between *formal learning* in educational institutions, which generally follows a course structure or curriculum and aims to provide learners with some form of degree or certification, and *non-formal learning*, referring to courses or programmes that individuals complete out of personal interest and which do not usually involve testing or certification (e.g. most online language learning apps).
3. The 'locus of control' denotes who is in charge of the learning, for example in terms of content, timing, etc. Terms such as *self-directed*, *independent*, or *autonomous* emphasise that the process is directed by learners themselves² (internal locus of control), rather than externally by teachers or instructors (external locus of control).
4. Finally, 'pedagogy' is concerned with the degree to which individuals are 'being taught' and 'have the intention of language learning' (Benson, 2011a, p. 11). In *self-instructed learning*, for example, individuals carry out activities designed for language learning for this explicit purpose. By contrast, *naturalistic learning* (in Benson's terms) occurs when individuals use their second language in activities not designed for learning, and without the explicit purpose of developing their language abilities.

² In line with the Oxford English Dictionary's guidelines, *they/them* will be used in this thesis as an inclusive (gender-neutral, non-binary) alternative to gendered singular third person pronouns (*he/him*, *she/her*).

While there is some conceptual overlap between Benson's four dimensions (e.g. most *formal learning* physically takes place inside of classrooms), this framework provides a way to define *informal second language learning* (ISLL), a form of language acquisition which is thought to result from activities such as reading, listening to music, watching films or television series, playing video games, or social networking in one's second language (L2).

These activities (henceforth *informal second language practices* or ISLPs) take place outside of educational institutions (*location*); are independent of formal frameworks for learning or testing (*formality*); and directed by the learners themselves (*locus of control*). Furthermore, they are not designed for language learning, and researchers seem to agree that individuals generally participate in them out of a genuine and personal interest (e.g. in a specific TV series, musician, video game, etc.) rather than for the purpose of language learning (*pedagogy*) (e.g. Olsson, 2011; Sockett, 2014b; Sundqvist, 2009, 2011; Sundqvist & Olin-Schneller, 2013). In fact, to emphasise that meaningful language use, rather than language learning, is the main focus for these individuals, Sockett (2014b, p. 14) has proposed that the term 'language user' should be used in place of 'language learner' in the context of ISLL (as previously in research on task-based learning and the use of English as a lingua franca). Similarly, Benson and Chik (2010) have concluded that the key feature of informal L2 learning is the 'situated and purposeful character' of the language use: the 'focus is on doing something with English rather than directly learning the language itself' (p. 68).

These quotes illustrate how research on ISLL can be situated among broader theories of language learning and thus has the potential to inform research on other areas within SLA: For example, there are parallels to usage-based models of learning

(e.g. Cadierno & Eskildsen, 2015; Ellis, 2017), which posit that learners first acquire language patterns or constructions which they frequently encounter and which are salient in the input, emphasising the importance of extensive L2 exposure. Similarly, informal second language practices can also be linked to *incidental learning*, referring to language acquisition which occurs as a ‘by-product’ of engaging in meaning-focused language activities (e.g. Hulstijn, 2001; Nation, 2001). This is not to say that individuals may not be aware of the potential of informal activities to support their language development or that they may not consciously attempt to learn a new phrase or grammatical feature they encounter during such activities. Indeed, the extent to which individuals pay conscious attention to linguistic forms during meaning-focused activities, and whether this is necessary or even beneficial to learning from them, has been a topic of much debate both in the field of ISLL (e.g. Cole, 2015; Olsson, 2011; Sockett, 2014b; Sundqvist, 2009) and in the broader SLA literature (cf. research drawing on the Noticing Hypothesis; Schmidt, 1990). Thus, exploring the role of focus on form in informal second language learning was one of the objectives of the current study.

A variety of terms have been proposed to refer to the specific type of informal learning described above, including: ‘naturalistic learning’ (e.g. Benson, 2011b; Benson & Reinders, 2011); ‘online informal learning of English’ (Sockett, 2014b; Toffoli & Sockett, 2010); ‘informal digital learning of English’ (J. S. Lee, 2017; J. S. Lee & Dressman, 2017); ‘fully autonomous self-directed learning’ (Cole, 2015; Cole & Vanderplank, 2016), ‘extramural English’ (Sundqvist, 2009, 2011); ‘unintentional learning of English’ (Forsman, 2004); and ‘language learning in the wild’ (Hannibal Jensen, 2019; Sauro & Zourou, 2019). It is worth noting that some of these terms

emphasise that a large part of informal learning takes place online and that many of them specifically reference the English language. While this reflects, in part, the general predominance of studies on English language learning in the SLA literature, the specific focus of ISLL research on the English language can be explained by the global prominence of Anglo-American media and culture and the wide-spread use of English as a lingua franca, especially on the internet (“Usage of content languages for websites,” 2019).

Despite the great number of terms already in use, the researcher decided to introduce the terms *informal second language learning* and *informal second language practices* in the current thesis. This is because the word ‘informal’ emphasises what the researcher perceives to be key aspects of this type of learning: that it takes place not just ‘beyond the classroom’ but also without the use of materials designed for the purpose of language learning (cf. Benson, 2011a). Additionally, referring to ‘second language learning’ in general, rather than the ‘learning of English’ leaves the possibility that some of the findings from the present thesis (although it, too, was concerned specifically with English language learning) may be transferable to the study of other languages. Whether this is indeed the case will, of course, remain subject to future research.

2.1.1 Describing Informal Second Language Practices

Some of the most detailed accounts of informal second language practices taking place online can be found in case studies from the field of *New Literacy Studies* (NLS), which is dedicated to exploring the communicative and literacy practices which emerge in new digital environments. NLS are highly relevant to research on informal second language learning because they have focused almost exclusively on out-of-

school contexts, and have consistently emphasised the value of informal practices as distinct from those associated with formal learning (Hull & Schultz, 2001). Studies have been conducted on a wide range of informal L2 practices, including the use of websites and forums, playing online role-playing and other video games, writing fan fiction, blogging, and social networking (for an overview, see Knobel & Lankshear, 2007).

While most NLS studies have been focused on English native speakers, two researchers have conducted in depth investigations of the online practices of second language learners: Black (e.g. 2005, 2006, 2009a, 2009b) conducted longitudinal studies of L2 fan fiction writers and Lam (e.g. 2000, 2004, 2006) ethnographically studied the online practices of young immigrant learners of English as a second language (ESL) in the USA. These in-depth case studies powerfully illustrate how the internet can provide opportunities for individuals with strong personal interests to use their second language to express their passion, enact their identities, and connect with like-minded people from across the world. They also show that even learners with only a basic level of proficiency can engage with vast amounts of personally relevant language input and communicate and collaborate socially with others in their second language. Furthermore, as the data for Lam's earliest study (Lam, 2000) was collected in 1997, it shows that online informal second language practices are not a particularly new or temporary phenomenon.

Notwithstanding, it must be questioned how common it is for language learners to use their L2 in their free time in the ways described in these studies. As Benson and Chik (2010) noted, such behaviours may appear common and widespread when studies explicitly focus on frequent internet users, but they may in fact prove to be exceptional when viewed in the context of general second language

learning. Therefore, this review will now turn to studies which have surveyed the informal second language practices of learner populations of various ages and in different educational and national contexts.

Sundqvist (2009, 2011; Sundqvist & Sylvén, 2014), for example, conducted detailed surveys on the out-of-school contact with English-language media of Swedish learners of English as a foreign language (EFL) of various ages. With the help of a semi-structured informal language learning diary (completed daily over the course of one week), the participants in her study reported how much time they spent with English books, newspapers and magazines, television, film, music and 'on the web'.

In her first study, Sundqvist (2009, 2011) found that that her 80 participants (aged 15-16) spent an average of 18.4 hours per week with English-language media, although there were vast individual differences (min = 0; max = 60). The activities on which her participants spent most time were, in order, listening to music, playing video games, and watching television and films, followed by surfing the internet and reading books, newspapers, or magazines. Sundqvist also found some gender differences in informal second language practices: The boys in her study reported spending significantly more time playing video games than the girls and slightly more time on English informal activities overall ($M = 21$ hours/week for boys vs. 16.4 for girls), although the latter difference was not statistically significant.

Sundqvist and Sylvén (2014) administered the same language diary to a group of 76 younger learners (10-11 years old). These participants spent substantially less time with English-language media per week than the older students, on average of 7.2 hours per week. Again, boys tended to have more contact with English than girls, mostly due to the time spent playing video games (in this study, the gender difference

was statistically significant). The activities on which these participants spent most time were, in order, watching television, listening to music, playing video games, watching films, surfing the internet, and finally reading. It is unclear whether the fact that television was more popular in this than in Sundqvist's previous study (2009, 2011) points to differences in informal practices among different age groups, or whether it is a function of the time at which the data were gathered (cf. studies by Toffoli and Sockett described below, which recorded significant gains in the popularity of television series among comparable samples between 2009 and 2012).

Kuppens (2007, 2010) conducted a study of English-language media consumption with a much larger sample of 361 Belgian students (aged 11-12), who had not yet received any English instruction at school. Even so, self-report questionnaires revealed that the participants had regular informal contact with English: 91% of them listened to English-language music at least three times per week, 57% watched English films or television series (mostly subtitled in Flemish), 47% browsed websites, and 37% played computer games in English. Like Sundqvist and Sylvén (2014), Kuppens also found that the boys played video games significantly more often, and also visited more websites in English than the girls.

Ten years later, De Wilde, Brysbaert and Eyckmans (2019) replicated Kuppens' (2007, 2010) study with an even larger sample ($n = 780$) of Belgian pupils of similar age (10–12 years old). The findings indicate that English-language activities had become even more popular in this learner population over time: 97% of the surveyed students now listened to English music on a daily basis, 80% watched English-language films or TV series subtitled in their L1, 78% used social media and 75% played video games in English. Whereas only 16% of the participants read books

in English every day, activities which involved speaking in English were surprisingly popular among the surveyed pupils: 45% reported speaking at least a few words or phrases in English every day, with friends and family, while gaming, or using social media.

In France, Toffoli and Sockett conducted large-scale surveys to study the *online* second language practices of university-level, non-specialist³ EFL learners in 2009 (n = 222; reported in Toffoli & Sockett, 2010) and 2013 (n = 225; reported in Sockett, 2014c). In both surveys, more than 95% of the respondents reported that they had 'regular' contact with English through the internet at least once per week. As in Sundqvist's (2009, 2011; Sundqvist & Sylvén, 2014) studies, listening to music and watching television series and films were by far the most common activities in both surveys. However, whereas watching TV series at least once per week was still much less common than listening to music in 2009, this gap seemed to have closed in 2012, suggesting a substantial gain in popularity of TV series during this time. Altogether, the participants engaged in receptive activities (listening or reading) more often than productive communicative activities (interacting, speaking or writing). Sockett (2014d) suggests that this might be because very few of the participants had personal contacts with whom they could communicate in English and even fewer belonged to English-speaking interest-based virtual communities.

The studies introduced above constitute some of the seminal research on informal second language learning, which emerged as a field of enquiry simultaneously, but

³ English language and literature were not their main subjects, but they did take English classes to complete the minimal language requirement for their degrees.

relatively separately, in several places around the same time. These studies provided initial evidence to show that informal English language practices (in which individuals engage outside of school, on- and offline) are generally wide-spread among L2 learners of various ages. That is, they do not seem to be limited to a small number of individuals with unusually high levels of interest in English and/or online media and culture, although there does seem to be considerable individual variation.

Since the early research discussed above, a great number of further studies have been conducted on the informal English language practices of a great variety of learner populations including, for example, young learners in Denmark (Hannibal Jensen, 2019); secondary-school learners in China (Lai et al., 2015), Greece (Lyriqkou, 2016, 2018), Sweden (Olsson, 2011; Olsson & Sylvén, 2015), and the Netherlands (M. Verspoor, De Bot, & van Rein, 2011); as well as university students in Brazil (Cole, 2015), France and Germany (Kusyk, 2017b), Hong Kong (Lai, Hu, & Lyu, 2018), Indonesia (2019), Slovenia (Jurkovič, 2019), and South Korea (J. S. Lee, 2017, 2019; J. S. Lee & Dressman, 2017; J. S. Lee & Lee, 2019). These studies have found that, even from a young age (e.g. De Wilde et al., 2019; Hannibal Jensen, 2019; Kuppens, 2007, 2010), learners engage widely in English-language activities, predominantly online, for an average of up to two hours per day (e.g. Cole, 2015; Lai et al., 2015; J. S. Lee, 2019; Olsson, 2011). English-language music, films, television programmes, video games, and websites were among the most popular media across all of these studies. Thus, it seems that ISLPs largely involve listening, while reading and writing are comparatively rarer forms of informal language contact. This is perhaps surprising, seeing as many of these activities take place on the internet, which is traditionally viewed as ‘being primarily a source of written information’ (Toffoli & Sockett, 2010, para. 17).

The conclusions that can be drawn from these studies are limited by the fact that all of them involved learners of English, which (as noted before) can be attributed to the prominence of this language within L2 teaching and research, as well as its status as a lingua franca in global off- and online discourse (Seidlhofer, 2005). There is evidence, however, that some learners of languages other than English also come into contact with their L2 in similar ways. Bailly (2011), for example, conducted interviews with more than 50 students in a school-supported autonomous language learning programme in France. The majority of interviewees reported regularly engaging in ISLPs which were strongly linked to their personal interests (listening to music, watching sports, films, or television, reading or playing games) in a variety of L2s, including Turkish, Italian, and Dutch. In Sundqvist and Sylvén's study (2014), some of the participants (albeit a 'negligible' number, according to the authors, p. 10) also reported engaging in informal activities in non-English languages, such as reading manga in Japanese, playing video games in German, or watching television and films and reading in Norwegian. Nevertheless, more systematic studies like those reviewed above are needed to determine whether ISLPs are as common among learners of languages other than English.

2.1.2 Identifying Outcomes of Informal Second Language Learning

In the previous section, the focus was on exploring the how common it is for second language learners, especially learners of English, to engage ISLPs. This section focuses on research that has investigated whether such practices are related to changes in language proficiency. Many such studies have been conducted which focus on the outcomes of one particular informal practice: For example, Kusyk and Sockett (2012;

Sockett, 2014a) demonstrated that French university students who spent more time watching English-language TV series scored higher on tests of receptive and productive vocabulary knowledge. Similarly, Sundqvist and Sylvén (2014) and Sundqvist and Wikström (2015) showed that, among Swedish primary and secondary school pupils, frequent video gaming was associated with increased receptive and productive vocabulary knowledge. However, as this thesis is concerned with ISLPs more generally, rather than one particular medium or activity, the discussion in this section focuses on the findings of studies which have investigated the relationships between informal L2 practices in general and language learning outcomes.

Correlational studies

Most ISLL studies to date have employed correlational methods to investigate the relationships between various measures of informal second language practices (e.g. the time individuals spend engaging in them, or how frequently they do so) and scores on various proficiency tests. In this way, significant, positive relationships have been discovered between ISLPs and vocabulary knowledge (De Wilde et al., 2019; Kuppens, 2007, 2010; Sundqvist, 2009, 2011; M. Verspoor et al., 2011); spoken (Lyriqkou, 2016; Sundqvist, 2009, 2011) and written language proficiency (Olsson, 2011; M. Verspoor et al., 2011); listening and reading comprehension (De Wilde et al., 2019); as well as self-reported language proficiency (Jurkovič, 2019).

Yet, links between ISLPs and proficiency measures were not confirmed in all studies: Lai, Zhu, and Gong (2015) and Lee (2017, 2019; J. S. Lee & Dressman, 2017) found no significant correlations between the time learners spent on informal L2 practices (quantity) and various measures of proficiency, although they *did* find significant links between language skills and the number of different activities in

which students engaged (variety). These studies are discussed in more detail in Section 2.2.2. In attempting to explain why his findings contradicted those of other ISLL studies, Lee (2019) speculated that the ‘quantity [of informal language contact] may produce different learning outcomes depending on the students’ learning contexts’ (p. 122). This seemed to occur in a study by Cole (2015), in which the quantity and variety of ISLPs correlated significantly with proficiency for ‘classroom-trained’ students, but not for ‘fully autonomous learners’ who were no longer receiving formal language instruction.

Sundqvist’s (2009, 2011) study of 15 to 16-year-old Swedish EFL learners (introduced in Section 2.1.1) also suggested that the relationship between ISLPs and language proficiency may vary between different groups of students. Across the whole sample, the amount of time the students spent on informal activities correlated moderately with both oral proficiency (evaluated through a five-part speaking test; $r_s = .31, p < .01$) and vocabulary knowledge (measured via Laufer and Nation’s [1999] Productive Vocabulary Levels Test and Nation’s [2001] Receptive Vocabulary Levels Test; $r_s = .36, p < .01$). However, the same analysis conducted separately for male and female participants indicated a strong gender difference: For the boys, informal L2 contact correlated strongly with vocabulary knowledge ($r_s = .59, p < .01$) and, to a slightly lesser extent, with oral proficiency ($r_s = .52, p < .01$). That is, 35% of the variance in the boys’ vocabulary scores and 27% of the variance in their oral proficiency could be attributed to the quantity of informal second language practices. By contrast, the same correlations were non-significant for the girls ($r_s = .01$ for vocabulary and $r_s = .12$ for oral proficiency, both $p > .05$).

Sundqvist’s (2011) attempt at explaining these gender differences references the types of resources activities in which the boys and girls in her study most

frequently engaged, and how these related to proficiency: Backward linear regression analyses revealed that playing video games, surfing the web, and reading books, newspapers, and magazines (what Sundqvist [2011, p. 114] calls 'active' or 'productive' activities; see also Section 2.2.2) were relatively more important contributors to variations in the oral proficiency and vocabulary scores than listening to music and watching films or television ('passive' or 'receptive' activities). Thus, the finding that the boys in this study spent significantly more time on 'productive' activities (particularly video games), which were more strongly correlated to language proficiency, might help to explain the aforementioned gender differences. Regrettably, Sundqvist did not report whether girls who *did* engage in these activities also demonstrated comparably higher levels of language proficiency than those who did not. Altogether, further evidence is needed to confirm the presence of gender effects in informal second language practices and learning.

It must be kept in mind that even though a majority of prior studies have found positive associations between ISLPs (quantity, frequency, and/or variety) and language proficiency, the direction of these relationships cannot be determined based on correlations. It is unclear whether informal activities positively affect learners' developing language skills (as is assumed by the overwhelming majority of the researchers cited above) or whether more proficient students are generally more able and willing to engage in ISLPs. The relationship between informal practices and proficiency may also be bidirectional, that is, the two factors may be mutually reinforcing over time. Thus, the next section focuses on studies which have employed alternative methods to establish links between ISLPs and proficiency, which in some ways address the issue of directionality.

Alternative approaches

Two of the studies previously introduced in Section 2.1.1 (De Wilde et al., 2019; Kuppens, 2007, 2010) focused on the informal practices of young ESL learners (10–12 years old) in Belgium. By studying learners who had not yet received any formal English language instruction (and excluding students who spoke English at home or had previously lived in an English-speaking country), this research provides some evidence for a directional relationship between ISLPs and language skills.

In addition to asking her 361 participants how frequently they consumed various types of English-language media, Kuppens (2007, 2010) also tested their ability to orally translate a small number of everyday phrases from English to Dutch and vice versa.⁴ A multivariate analysis of variance (MANOVA) of engagement in various informal activities and translation scores indicated that watching English subtitled films or television more often (6–7 times per week, vs. 3–5 times or less than 3 times) was associated with significantly better skills both in the L1-L2 and L2-L1 direction ($p < .001, \eta^2 = .03$, indicating a small effect). This link was stronger among the female than the male participants (significant interaction effect at $p < .01$, partial $\eta^2 = .01$), which Kuppens took to indicate that ‘boys and girls [...] watch English subtitled television and movies in a different way’ (2010, p. 78). The analysis also indicated a significant but weaker effect of playing computer games on translation skills ($p < .05$, partial $\eta^2 = .01$).

In their follow-up to Kuppen’s (2007, 2010) study, De Wilde et al. (2019) collected data about the average daily duration of 780 students’ informal contact with English

⁴ In language learning research, L2–L1 and L1–L2 translation skills often serve as proxy measures for, respectively, receptive and productive language skills.

and also administered the Peabody Picture Vocabulary Test (Dunn & Dunn, 2007, a standardised measure of receptive vocabulary size for young learners) and the Cambridge English Test for Young Learners (“Cambridge English Qualifications: A2 Flyers,” 2019), which tests listening and reading comprehension, writing, and speaking. Significant correlations emerged between the proficiency measures and all of the tested informal activities (watching TV and films with or without L1/L2 subtitles, listening to music, reading books, playing video games, using social media, and speaking), although some of them were very weak. *Using social media* was most strongly correlated to the different proficiency scores (Pearson’s r between .39 and .44, all $p < .001$), whereas the links were weaker for *watching TV or films with L1 or L2 subtitles* (some significant correlations, r between .08 and .12, $p < .05$). This latter result is in direct contrast to Kuppens’ (2007, 2010) study, which found that watching TV and films was more strongly linked to English translation skills than any of the other informal activities under investigation.

De Wilde and colleagues (2019) also carried out multiple regression analyses to directly *compare* the effects of the different activities on the different types of proficiency. Similar to Sundqvist (2009, 2011), they concluded that productive activities contributed relatively more to these young learners’ English language skills than receptive activities: Whereas using social media and, to a lesser extent, speaking in English predicted all five proficiency measures (respectively, β between .29–.36 and between .13–.21, all $p < .001$), playing games had significant effects only on receptive vocabulary size ($\beta = .16, p < .001$) and speaking skills ($\beta = .12, p < .01$). Interestingly, listening to music in English emerged as a very weak but significant *negative* predictor of reading comprehension and writing ($\beta = -.10, p < .05$) and receptive vocabulary ($\beta = -.08, p < .05$). The researchers tried to explain this

finding by arguing that listening to music, which does not necessarily lead to understanding or learning the language involved, takes away time from other, more effective informal learning activities. Together, all of the significant predictors explained between 18 and 23% of the variance in the different proficiency scores.

The studies conducted by Kuppens (2007, 2010) and De Wilde and associates (2019) provide some evidence in favour of a directional relationship between informal second language practices and L2 proficiency, although they do not rule out that there also might be an effect of prior language proficiency on L2 media use. This cautious interpretation is supported by the fact that Kuppens (2007, 2010) chose only to test the correlations between translation abilities and frequency of exposure to three types of media (subtitled films and TV programmes, music, and computer games) while excluding sources such as non-subtitled television and films, radio, and websites, which she felt required ‘a certain English proficiency’ to be ‘used meaningfully’ by the pupils (2010, p. 72). However, the exact criteria by which she judged that the latter activities required greater prior proficiency, or how she defined ‘meaningful use’, are unclear. As will be discussed later in this chapter (Section 2.2.2), there is a lack of research which looks at the different ways in which students interact with various informal second language practices and any effects on the associated language development.

Verspoor, De Bot and van Rein (2011) and Cole (2015; Cole & Vanderplank, 2016) studied the effects of ISLPs on developing language proficiency by making direct comparisons between different groups of learners. The former study investigated the

language skills of students with and without access to informal resources, while the latter focused on ISLL in students with and without concurrent formal instruction.

Verspoor and colleagues (2011) employed a two-by-two-by-two ($2 \times 2 \times 2$) design, comparing (a) students in their first or third year of secondary schools (ages 13 and 15), who were (b) enrolled either in English-Dutch bilingual schools or Dutch monolingual schools, where they received only a few hours of English instruction per week and who (c) either had regular access to English-language popular media (e.g. television, music, and the internet) or whose access to these resources was severely limited for religious reasons (*high vs. low informal access*). All participants completed two English-language proficiency tests at three points during one school year: a receptive vocabulary inventory ('I do/do not know this word') and a narrative writing task, which was scored based on a range of criteria including lexical variation, syntactic complexity, and linguistic accuracy. The group mean scores were compared statistically across grade levels, media access, and enrolment in mono-/bilingual education. It should be noted that some caution needs to be exercised when interpreting these results due to the uneven sizes of the subsamples (there were only 26–31 people in each *low informal access* groups, vs. 83–184 students in the *high access* groups).

As expected, the students who were allowed to access popular media had significantly more out-of-school contact with English than their limited-access peers ($p < .001$, partial $\eta^2 = .14$, large effect). In terms of the effects of media access on proficiency, the data indicated that, at the beginning of secondary school, there were no significant differences in proficiency between the high and low informal access groups. Nevertheless, a lack of informal L2 contact did seem to have a significant long-term effect on the students' English language development. These effects occurred at

different times for the different types of knowledge tested: There were significant differences between the first-year low- and high-access groups on the second writing test ($p < .01$, partial $\eta^2 = .08$, medium effect), which further increased in the third test ($p < .001$, partial $\eta^2 = .17$, large effect). In the third-year group, the high informal access group outperformed the low access group on all three receptive vocabulary and writing tests (all $p < .001$) and these differences were even more pronounced (partial η^2 between .28 and .40, indicating very large effects).

In fact, the frequency of engagement with popular media correlated significantly with the productive writing skills of students in all age groups and forms of education (Pearson's r between .23 and .41, all $p < .01$) and with receptive vocabulary for the older students (Pearson's r between .25 and .48, all $p < .01$). The authors concluded that these correlations indicate the impact of popular media on proficiency over time, first affecting productive (writing) skills and later receptive skills (vocabulary). This is in contrast to prior studies which have generally found equally strong or even stronger links between ISLPS and vocabulary knowledge than with productive language skills (e.g. De Wilde et al., 2019; Kuppens, 2007, 2010; Sundqvist, 2009, 2011).

In terms of the interaction effects between formal language instruction and informal language contact, the authors concluded that bilingual education successfully 'compensated' for a lack of media access: The low informal access students in bilingual schools scored as well on the proficiency tests as their peers in monolingual programs, but still below the level of the high-access students in bilingual schools. Altogether, the findings of this study strongly support the hypothesis that ISLPs can promote the development of L2 proficiency, even above and beyond what is achieved through intensive formal language contact.

Cole (2015) also investigated the directional relationship between ISLPs and L2 proficiency by comparing two distinct groups of learners in Brazil: ‘fully autonomous self-instructed learners’ (FASILs), who engaged exclusively in informal learning of English, and classroom-trained learners (CTLs), who attended English classes in prestigious private language schools. Cole recruited 84 non-specialist students (34 FASILs and 50 CTLs) in post-secondary education (ages 18 to 24) who considered themselves upper-intermediate to advanced learners of English. The participants completed seven proficiency tests, including reading comprehension, narrative writing, lexical recognition, grammaticality judgment, and error correction tasks. All tests were timed, as the aim was to measure the learners’ automatized, ‘deep’ knowledge of lexical and grammatical structures, as well as their fluency in reading and writing.

Statistical comparisons of the two groups’ mean scores showed that the FASILs performed significantly better than the CTLs on all seven tests (all $p < .01$, Cohen’s d between .76 and 1.58, indicating large to very large effects). Cole also reported differences in the types of errors made by the two groups of learners, observing that the FASILs, more so than the CTLs, seemed to have overcome cross-linguistic transfer errors common among Brazilian learners almost completely and acquired a nuanced, native-level command of English.

Altogether, these findings suggest that informal second language practices can support the development of L2 proficiency. Cole and Vanderplank (2016) also point out that these results challenge the assumption (wide-spread in the field of SLA) that late-onset language learners cannot reach high levels of proficiency without formal language instruction. This conclusion must be approached with some caution, however, as the ‘fully autonomous self-instructed learners’ were not completely

without formal training: All of them had studied English as part of the general Brazilian primary and secondary school curriculum, where they had presumably achieved at least basic L2 proficiency. Furthermore, the participants who were categorised as FASILs were allowed to have been enrolled in private language schools or to have made use of other formal learning methods (e.g. distance learning courses or private tutoring) for up to one year. Thus, on the basis of Cole's (2015) study it cannot be completely ruled out that it is the level of L2 proficiency which affects ISLPs rather than vice versa. That is, it is unclear whether the FASILs in this study were more proficient in English because they had engaged in more ISLPs, or whether those individuals who had previously gained a higher proficiency in English (i.e. from mandatory formal English education in school) chose to continue to learn the language informally instead of enrolling in private language schools. In fact, the latter seems to apply to at least some of the FASILs, who (in the semi-structured interviews which were also part of Cole's study) expressed that they disliked formal learning because they considered it inauthentic and inaccessible, and therefore less able to support their L2 development.

For all of these reasons, it cannot be deduced that informal learning *alone* was responsible for the high levels of language proficiency that the FASILs demonstrated in Cole's study. In fact, an even more complex picture emerged from an analysis of the correlations between a combined proficiency measure and the quantity and variety of ISLPs (the average time learners spent on informal practices per week and the number of different activities in which they engaged 'regularly'). Both of these variables correlated significantly with proficiency in the sample of CTLs but not among the FASILs. Furthermore, when these variables were entered into a multiple regression model, together with other factors which showed significant correlations

with proficiency (including the total number or years spent learning English, and various motivational and attitudinal factors), the quantity and variety of ISLPs did not emerge as significant predictors. A binary variable representing whether participants were ‘classroom-trained’ or ‘fully autonomous’ did, however, significantly predict proficiency, as did their language learning motivation, to some extent. Cole therefore theorised that it was not only the informal language use which made the FASILs more proficient than the CTLs, but their different attitudes towards learning and using the language (more details in Section 2.3.2).

Finally, there are longitudinal studies of ISLL, which were conducted by Kusk (2017b) and Olsson and Sylvén (2015). Kusk (2017b) tracked the language development of three informal learners over a period of five months using measures of complexity, accuracy, and fluency. She observed great inter- and intraindividual variation over time, both in the individual learning trajectories and the informal practices themselves, which illustrates the complexity and non-linearity of ISLL. However, Kusk did not report extensively on the extent to which the participants’ informal practices and their linguistic skills co-evolved.

Meanwhile, Olsson and Sylvén (2015) conducted a three-year longitudinal study on informal second language practices and learning among 146 Swedish students in *content and language integrated learning* programs (CLIL, where English is the language of instruction in some or all classes) and 84 students at monolingual Swedish schools. Data on the participants’ ISLPs was collected in their first year of secondary school, through a survey which measured the frequency of students’ engagement in various informal practices and a language diary, in which pupils recorded the time they spent on such activities (quantity of ISLPs). In addition, Olsson

and Sylvén also tracked the development of the students' academic vocabulary over the course of the three years by assessing the occurrence of academic words in four argumentative and explanatory essay-writing tasks (two in the first and one in each of the subsequent years).

There were some significant differences between the CLIL and non-CLIL students' patterns of informal L2 contact. Most notably, the CLIL students spent on average more time on ISLPs per day than the children in monolingual schools ($p < .01, d = .49$, medium-sized effect), although there was substantial inter-individual variation within both groups. However, the study found no significant effects of the frequency and duration of informal language contact on the development of academic vocabulary in the two groups. Altogether, then, this study does not provide evidence for a positive effect of informal second language practices on the development of, specifically, academic vocabulary.

A major limitation of Olsson and Sylvén's (2015) research was that the participants' informal language contact was measured only at the beginning of the study. In fact, the frequency (but not the quantity) of ISLPs *did* correlate significantly with the vocabulary scores at the first point of data collection ($r = .25, p < .01$). This suggests that the study might have found significant links between ISLPs and proficiency if it had explored the relationships between repeated measures of informal L2 contact and proficiency over time.

For the most part, the studies reviewed in this section have found a positive relationship between informal second language practices and L2 proficiency. A few studies have even provided some evidence for a directional effect of informal L2 contact on language proficiency (Cole, 2015; Kuppens, 2007, 2010;

M. Verspoor et al., 2011), although further research is certainly needed to explore directionality in ISLL.

Nevertheless, the findings regarding the link between ISLPs and language skills have not been unanimous. Some studies found that the variety of informal activities was more strongly related to proficiency than the time spent on ISLPs (Lai et al., 2015; J. S. Lee, 2017, 2019; Olsson & Sylvén, 2015), whereas others discovered significant correlations only for particular sub-groups of learners (Cole, 2015; Sundqvist, 2009, 2011). These diverse findings, then, support the conclusion that the relationships between informal second language practices and L2 proficiency are complex.

Attempts to explain variations in results have generally been vague, with some researchers proposing that variations in learning outcomes might have occurred due to differences in learner motivation, or because different (groups of) learners may ‘interact with’ informal resources in different ways (e.g. Cole, 2015; Kuppens, 2010). Thus, further research is needed at this point which systematically captures the *nature* of learners’ informal second language practices (beyond the frequency, quantity, and variety of activities) and investigates the relationship between these factors, motivation, and learning outcomes over time. With the current study, the researcher aimed to address this gap in the ISLL literature.

The following section introduces the concept of *engagement*, which was chosen as a framework for conceptualising the *nature of informal second language practices*, and therefore gaining a better understanding of what might make some learners more likely to participate in, and benefit from, informal L2 activities.

2.2 Engagement

2.2.1 Theoretical Background

Although it is a relatively new concept in SLA research (e.g. Akbari, Ahmad, Simons, & Pilot, 2016; Asoodari, Atai, Vaezi, & Marandi, 2014; C.-C. Liu, Wang, & Tai, 2016), *engagement* is a term widely used and readily understood among teachers and has been studied in Educational Psychology for decades. It is broadly accepted that there is a relationship between engagement and academic performance, which has been demonstrated extensively in research (concluded, for example, by Boekaerts, 2016; Finn & Zimmer, 2012; Fredericks, Blumenfeld, & Paris, 2004). Nevertheless, there is a lack of definitional clarity around the concept of engagement — owing to the use of widely varying terminology, working definitions, and research tools — which makes it difficult to consolidate empirical findings. The issue is further complicated by the fact that engagement is often treated as an umbrella term, under which other related concepts (e.g. motivation, anxiety, boredom) are included by some researchers but not others (Finn & Zimmer, 2012).

Based on their own review of the engagement literature, Christenson, Reschly, and Wylie (in their *Handbook of Research on Student Engagement*; 2012) offered the following definition:

[E]ngagement refers to the student's active participation in academic and co-curricular or school-related activities, and commitment to educational goals and learning. Engaged students find learning meaningful and are invested in their learning and future. It is a multidimensional construct [...]. Student engagement drives learning; requires energy and effort; is affected by multiple contextual influences; and can be achieved for all learners. (p. 816-817)

This definition touches upon a few points where a consensus seems to have been reached in the engagement literature. These are addressed in the following paragraphs, together with short summaries of any associated issues that do persist.

1. Engagement at different levels

Students can be engaged at different ‘levels’, such as in their overall education, the school which they attend, particular subjects, or even in a single task. Most research to date has focused on school-level engagement, whereas studies at the subject- and task-level, as well as those investigating engagement in learning outside of the classroom, are relatively less common (Eccles & Wang, 2012; Reschly & Christenson, 2012).

2. Engagement as a multidimensional meta-construct

There is broad agreement that engagement comprises at least three different *dimensions* (or *aspects*): behavioural, cognitive, and affective engagement (Fredericks et al., 2004). Some researchers extend this tripartite framework to also include, for example, social or self-regulatory aspects (Boekaerts, 2016; Finn & Zimmer, 2012). Most scholars would agree that the different dimensions of engagement are dynamically interrelated, that is, they mutually influence one another (see contributions in Christenson et al., 2012).

Besides a disagreement over the precise number of engagement aspects, there is again a lack of conceptual clarity as these dimensions are operationalised in questionnaire items and measurement scales: ‘making an additional effort’, for example, arguably measures either cognitive or behavioural engagement, depending on whether it is considered an observable behaviour or learner-internal cognitive

process (Boekaerts, 2016). In fact, because cognitive and affective engagement are internal processes, they are difficult to measure and often operationalised via behaviours (Appleton et al., 2006; Finn & Zimmer, 2012), such as observable learning strategies. This further contributes to the blurring of boundaries between the different dimensions of engagement and impedes the synthesis of research findings.

3. Engagement drives learning

As stated previously, the relationship between engagement and academic performance is widely accepted. However, there is some difficulty in distinguishing engagement from another construct, which is similarly described as ‘driving’ learning (e.g. Dörnyei & Ryan, 2015): *motivation* (see Section 2.3). Most researchers have tried to delineate borders between motivation and engagement, treating them as separate but related constructs (Christenson et al., 2012). Often, motivation is described as an antecedent to engagement, wherein motivation refers to the underlying reasons for a given behaviour, and engagement is seen as ‘the visible manifestation of motivation’ (Skinner, Kindermann, Connell, & Wellborn, 2009). This approach, however, seems incompatible with the conceptualisation of engagement as multidimensional, since cognitive and affective engagement are not easily ‘visible’.

Instead of limiting engagement to the outwardly observable, it may be more appropriate to define it as ‘the observable and unobservable qualities of students’ interactions with learning activities’ (Wang et al., 2016, p. 17) and motivation as an underlying psychological process which ‘moves a person to make certain choices, to engage in action, to expend effort, and to persist in action’ (Dörnyei & Ushioda, 2011, p. 3). In other words, engagement can be seen as referring to the *nature* of individuals’ participation in learning activities (e.g. the time they spend on them,

which activities they choose, what they are feeling and thinking while engaging in them) and motivation as encompassing *why* they participate in this way (e.g. reasons for learning, how individuals relate to the language and the learning context, see Section 2.3).

4. Engagement is Contextually Dependent

Engagement is generally viewed as contextually dependent. For example, students' levels of engagement can vary depending on the particular subject, classroom practices, or whether they are learning independently or in a group (Boekaerts, 2016). Additionally, some researchers treat engagement as dynamic across time, too, while others view it as a relatively stable characteristic, which changes very little in the long term (Eccles, 2016). This is related to the *level* at which engagement is studied (cf. pt. 1 above): Task-level engagement is, by definition, situated and dynamic construct, whereas school-level engagement is often viewed as more stable. A contributing factor may be that state-like variables are easier to operationalise in research, requiring, theoretically, only a one-time measurement. However, if engagement is defined as contextually and temporally dependent, measures must be domain-specific and context-dependent and long-term studies and repeated measures across varying contexts are needed to formulate and test an overarching theory (see Lambert, Philp, & Nakamura, 2016 for a good example).

5. Engagement can be achieved by all learners

Perhaps the most important assumption underlying this area of research is that all individuals can become engaged in learning. Thus, the ultimate goal of many researchers and practitioners is the improvement of academic outcomes for all

students by promoting higher overall levels of engagement, or certain types of engagement over others. However, to design effective programmes to improve students' learning, we need a detailed understanding of which aspects of student engagement influence which aspects of learning and performance, and the extent to which these relationships may vary between learners and in different contexts.

Despite some of the theoretical issues around the definition and operationalisation of the construct, the researcher decided that engagement could serve as a suitable theoretical framework in the current study for investigating the nature of informal second language practices, or the 'qualities of students' interactions with [informal] activities' (cf. Wang et al., 2016, pt. 3 above). For this purpose, the three dimensions included in Fredericks et al.'s (2004) original tripartite framework were adopted with the following working definitions:

1. *Behavioural engagement*: the observable qualities of engagement in ISLPs, including the duration, frequency, and variety of informal activities
2. *Affective engagement*: learners' emotional states while engaging in ISLPs, including, for example, enjoyment, interest, and challenge-skill balance
3. *Cognitive engagement*: the extent to which learners are concentrated and focusing on the task at hand versus on extraneous stimuli

In addition, the researcher proposed to include a fourth dimension, to increase the applicability of the engagement construct to the study of language learning in particular:

4. *Linguistic engagement*: the extent to which learners are focused on processing linguistic features and improving their language skills (i.e. focus on form as opposed to focus on meaning)

Like cognitive engagement, this construct also describes the learner's mental state while engaging in informal activities. However, while cognitive engagement refers to the amount of mental effort that learners put into engaging with the content of these activities themselves, linguistic engagement more specifically encompasses effort dedicated to processing linguistic features encountered in the activity.

Through the addition of linguistic engagement, this framework also connects the current research with two broader areas of enquiry within SLA: First, since linguistic engagement is used to operationalise the extent to which learners pay conscious attention to form in their informal second language activities, there is a strong link to the vast body of work on *incidental language learning* (see Section 2.1). Second, there is overlap with research on *language learning strategies*, that is, techniques which individuals employ in order to help them in, for example, comprehending or learning new linguistic features (Chamot, 2004; Cohen, 1998). Although various taxonomies of strategies cover both mental processes (e.g. inferring the meaning of a new word from context, creating a mental image to aid memorisation) as well as observable behaviours (e.g. looking up a word in the dictionary or taking notes) (see review in Grenfell & Macaro, 2007), it has been argued that the latter can be viewed as overt manifestations of underlying cognitive strategies (Macaro, 2006; Stevick, 1990). In the current theoretical framework, it is assumed that individuals who employ strategies in order to increase their comprehension, or to try and learn new linguistic features they encounter in their

informal second language practices, will report higher levels linguistic engagement (conscious focus on linguistic features), and thus, indirectly, higher levels of cognitive engagement (paying closer attention to the activity at hand).

In addition to comprising the aforementioned four dimensions, engagement is viewed in the current study as highly situated, or task dependent. This in particular distinguishes engagement from *attitudes* (see Section 2.3.1), which similarly comprises affective and cognitive (and by some definition also behavioural) components (Jhangiani, Tarry, & Stangor, 2014). Whereas attitudes refer to learners' general beliefs about or feelings towards a particular subject (Tódor & Dégi, 2016) — in this case, the context for language learning (informal or formal) — engagement comprises participants' behaviour, thoughts, and feelings *while engaging in* a particular activity.

The following section gives an overview of how the nature of learners' engagement in informal second language practices has been operationalised in previous studies and what they have found with regard to the links between different aspects of engagement and L2 proficiency. Although these studies did not adopt the same theoretical framework, wherever possible, their results will be discussed in terms of the different dimensions of engagement as defined above to facilitate comparability with the results of the current study.

2.2.2 Engagement in Informal Second Language Learning

Behavioural engagement: Quantity and variety of informal practices

As stated previously (section 2.1.1), prior research has focused mainly on measuring the frequency with which learners engage in informal second language practices (e.g. Jurkovič, 2019; Kuppens, 2007, 2010; Lyrigkou, 2016; Olsson & Sylvén, 2015; Sockett, 2014d; Toffoli & Sockett, 2010) and/or how much time they spend on them (e.g. Cole, 2015; De Wilde et al., 2019; Olsson, 2011; Sundqvist, 2009, 2011; Sundqvist & Sylvén, 2014) — both aspects of behavioural engagement. In this section, two studies will be discussed in more detail, which have considered an additional aspect of behavioural engagement: the variety of activities in which learners engage.

Lai, Zhu and Gong (2015, see also Section 2.1.2) investigated the of out-of-class L2 activities of 82 EFL learners in China (14 years old) using questionnaires and focus group interviews. 'Out-of-class learning' in this study included both informal activities (e.g. watching television, reading, or listening to music) and *non-formal* activities (e.g. attending tutorial classes and engaging in form-focused language practice). Despite this broader scope, the results are clearly highly relevant to ISLL research.

The quantitative analyses revealed significant correlations between the variety of their participants' out-of-class learning (the number of different types of activities in which they engaged regularly) and the participants' English grades (Pearson's $r = .25, p < .05$), their confidence in themselves as learners of English ($r = .42, p < .01$), and their enjoyment of learning ($r = .45, p < .01$). However, the *quantity* of out-of-class activities (the average time spent on them per day), did not correlate significantly with any of the criterion measures. Meanwhile, in analysing their qualitative data, Lai et al. observed that the respondents with the highest levels of

proficiency, confidence, and enjoyment mostly engaged in informal, rather than non-formal activities.

In a similarly designed project, Lee (2017, 2019; J. S. Lee & Dressman, 2017) also studied the links between the quantity and variety of out-of-class activities and various measures of language knowledge. Data were collected from 94 university-level learners of English in South Korea, who completed the receptive and productive Vocabulary Levels Test (Laufer & Nation, 1999; Nation, 2001) and a speaking test in addition to providing a retrospective self-report of their out-of-class activities in the six months prior to the study.

The variety of the activities in which the learners engaged significantly correlated to receptive vocabulary size ($r = .27, p < .05$; Lee, 2017), the productive vocabulary scores (r between .46 and .54⁵, $p < .01$; Lee, 2017, 2019; Lee & Dressman, 2017), performance on the speaking test (both $r = .56, p < .01$; Lee, 2019; Lee & Dressman, 2017), and standardised university entrance English exam scores (TOEIC; $r = .47, p < .01$; Lee, 2019). By contrast, the time that the participants spent on out-of-school practices did not significantly correlate to any of these measures. Both the quantity and variety of informal activities did, however, share significant correlations with the learners' enjoyment of and confidence in language learning (J. S. Lee, 2019, see also Section 2.3.2).

In trying to explain why the quantity of out-of-school practices was significantly correlated to language skills in many prior studies (see Section 2.1.2.) but not his own, Lee (2019) speculated that 'quantity may produce different learning

⁵ Different learners were excluded from various analyses in each of the publications, resulting in slight differences in the reported statistics.

outcomes depending on the students' learning contexts' (p. 122). In other words, he postulated that there may be differences in the ways in which engagement in out-of-class practices contributes to language development for East Asian students (such as those in his research, as well as Lai et al., 2015) versus learners in other (mostly European) contexts. At the very least, both of the studies discussed above emphasise that measuring engagement in ISLPs in various ways is necessary to expand our understanding of the circumstances in which these activities may contribute to language development.

Linguistic and Cognitive Engagement: Attention to Form and Meaning

Beyond whether or not learners engage in certain informal practices, and how much time they spend on them, the kind of attention they pay to these activities has also been a topic of discussion among researchers in ISLL. In particular, a number of researchers have proposed that a focus on form (i.e. linguistic engagement) during ISLPs promotes greater language development. For example, Sundqvist and Sylvén (2016), after reviewing previous findings regarding the types of ISLPs that are most popular among learners and their correlations with proficiency, proposed a distinction between 'active' and 'passive' practices (also used in Sundqvist, 2009, 2011). *Passive* activities, according to Sundqvist and Sylvén (2016), largely involve receptive language use and include, for example, listening to music and watching TV or films, whereas *active* practices are those in which learners 'need to be active and rely heavily on their own L2 English language' (p.188). This includes mostly productive language activities (e.g. speaking, writing), but also some which involve receptive language use, like reading and playing video games. By labelling certain informal practices as 'active' or 'passive', the researchers implicitly judge how

mentally engaged learners are in certain activities. This distinction seems to be based purely on their own experiences and theories, as they did not measure learners' attention in various types of informal activities in any of their prior studies (Sundqvist, 2009, 2011; Sundqvist & Sylvén, 2014). According to Sundqvist and Sylvén (2016), while it 'takes more effort to engage' in active practices, they are also 'relatively more important' for gaining L2 proficiency (p. 139). That is, the researchers proposed that a certain level of focus on language is beneficial, if not necessary, for successful informal second language learning.

A similar position is shared by Cole (2015, for a summary of the design, see Section 2.1.2). He observed that the 'fully autonomous self-instructed learners' (FASILs) in his study (who, at the time, were using their L2 [English] frequently, but were no longer receiving formal instruction) reported investing more 'effort' into ISLPs than 'classroom-trained learners' (CTLs). This means that the FASILs more frequently talked about intentional learning behaviours, such as using subtitles when watching L2 videos or googling unknown words or phrases. Cole suggested that the higher use of such comprehension strategies among FASILs could explain why their language proficiency was significantly above that of the CTLs (demonstrated in seven tests they completed as part of the study, see Section 2.1.2). Furthermore, Cole (2015) argued that the different levels of linguistic engagement among FASILs and CTLs may, in turn, be explained by the learners' different *motivation* and *attitudes* towards ISLL in general.

In particular, Cole observed that the CTLs considered their language classes the most valuable learning resources at their disposal and may thus have perceived ISLPs merely as leisure activities — therefore investing less 'mental effort' (p. 339)

while engaging in them, and thus learning less through them. On the other hand, the FASILs seemed to consider informal activities their most important resources for learning *as well as* enjoyment, which Cole speculates may be what led them to invest more cognitive effort and make greater gains in proficiency (cf. Salomon, 1984). In other words, Cole proposes that individuals' attitudes towards informal second language learning affect the amount of effort they invest into decoding the language they encounter in ISLPs (linguistic engagement) and, in turn, their language development. While this is a compelling theory, it must be noted that it was formulated based on participants' self-reports in response to rather general interview questions; like Sundqvist and Sylvén (Sundqvist, 2009, 2011; Sundqvist & Sylvén, 2016), Cole did not include a focused measure of *linguistic engagement* in his study.

Lyriqkou (2016, 2018), on the other hand, directly examined the 'amount of cognitive effort' which her participants, 76 Greek secondary school EFL learners (ages 13–16), invested in understanding and learning language features they encountered while engaging with informal practices. Using self-rating scales, her participants reported how likely they were to, for example, try to guess the meaning of unknown words or phrases from context or search for translations in their first language. No significant correlation was found, however, between the level of effort and the participants' scores on a test measuring spoken language proficiency. Therein, these findings contradict Cole's (2015) and Sundqvist and Sylvén's (2016) suggestion that linguistic engagement in ISLPs would correlate positively with language skills.

Similarly, Lai, Zhu and Gong (2015) do not agree that a focus on form is essential to successful informal L2 learning. In semi-structured interviews with secondary school pupils in China, they observed that the less proficient learners

tended to report focusing on linguistic forms more often than learners with higher proficiency, even in so-called ‘meaning-focused activities’ (i.e. *informal practices*, as opposed to *non-formal* activities such as voluntary grammar exercises or tutorial classes). Based on this, the researchers suggest that a focus on *meaning* might be particularly important to successful language development, enjoyment, and interest in out-of-class learning because it complements the much more form-focused approach of in-classroom learning.

In summary, research on the nature of learners’ engagement in informal second language practices has largely focused on behavioural engagement (the quantity and diversity of informal activities). Studies have suggested that, in some populations, the variety of activities in which learners engage may be relatively more important for ISLL than the amount of time spent on them (Lai et al., 2015; J. S. Lee, 2017, 2019; J. S. Lee & Dressman, 2017). Meanwhile, even though several researchers have debated whether a focus on form in ISLPS is beneficial or even necessary for successful informal language learning (Cole, 2015; Lai et al., 2015; Sundqvist & Sylvén, 2016), there is a considerable lack of studies which have actually measured attention in ISLL (however, see Lyrigkou, 2016, 2018). Finally, it seems that studies which consider the role of emotions (*affective engagement*) in informal second language practices are altogether missing from the literature at this point. This thesis seeks to address these gaps in the ISLL literature by collecting data on learners’ behavioural, affective, cognitive, and linguistic engagement and modelling the relationship between these variables and changes in language proficiency over time.

2.3 Motivation and Attitudes

Motivation is regarded as one of the most important factors in successful second language learning (Dörnyei & Ryan, 2015) and it has been an area of focus in SLA research for more than fifty years. Studies have generally found positive links between language proficiency and high levels of motivation (e.g. Toni & Rostami, 2012; Vandergrift, 2005; Zhang & Xiao, 2006) although the direction of this relationship, or whether these variables are reciprocally related, remains a topic of much discussion (Gardner, 2000; X. Liu, 2016).

Concern with motivation has primarily grown out of an interest in the role of individual differences in language learning. The concept has its roots in Psychology, where it has been defined as a mental process that ‘serve[s] to activate or energize behavior and give it direction’ (Huitt, 2011, para. 1). In SLA, similar definitions have been adopted. Dörnyei and Ushioda, for example, described motivation as ‘what moves a person to make certain choices, to engage in action, to expend effort, and to persist in action’ (2011, p. 3).

As stated in the introduction (Chapter 1), one of the goals of the current thesis was to understand students’ *perceptions of language learning* affect, and are affected by, engagement in informal second language practices. Theories of motivation in SLA have traditionally addressed one of more of three aspects of learner perceptions (Dörnyei, 2001; Kormos, Kiddle, & Csizér, 2011):

1. *Goals*: students’ perceptions of the language itself and why it would be useful for them to know

2. *Attitudes*: perceptions of the learning context
3. *Conceptions of the self and identity*: perceptions of themselves as language learners and users.

In order to gain a better understanding of these concepts, the following sections provide an introduction to how they have been conceptualised in three leading theories of motivation in SLA (as identified in reviews, e.g. by Al-Hoorie, 2017; Boo, Dörnyei, & Ryan, 2015; Dörnyei & Ushioda, 2011; Sugita McEown, Noels, & Everhart Chaffee, 2014): (1) the Socio-Educational Model; (2) Self-Determination Theory; and (3) the Second Language Motivational Self-System. The researcher decided to focus on these theories, which have primarily been developed in and applied to the study of formal language learning, in order to explore their applicability to the informal second language learning context. This will be followed by a review of prior ISLL studies which have also included measures of language learning motivation.

2.3.1 Theoretical Background

The Socio-Educational Model and Language Learning Orientations

The *Socio-Educational Model* (SE Model; Gardner, 1985) is one of the earliest and most influential theories of motivation in SLA (Dörnyei, 2005, p. 72). According to this model, there are three components to language learning orientations: a desire to learn a second language; being willing to expend effort towards this goal; and having positive attitudes towards the learning situation (Gardner, 2010). The desire to learn a second language is operationalised in the SE model via the concept of *integrativeness*, which itself encompasses an admiration for the target language

community; a desire to integrate into this culture (*integrative orientation*); and a general desire for second language learning (Masgoret & Gardner, 2003). The SE model was developed in the bilingual (and to some extent bicultural) Canadian context and the concept of integrativeness was founded on the idea that second language learning involves not just the acquisition of skill but also developing a new identity, as well as learning to interact in new social contexts.

Out of the many components of the SE model, the integrative orientation has become the most emphasised and well-researched (Dörnyei, 2003; Sugita McEown et al., 2014). Later on, Gardner and his colleagues introduced a second language learning orientation into the SE model, the *instrumental orientation*, which involves wanting to learn a second language for utilitarian reasons (*benefits*, such as becoming better-educated, obtaining a better job, or earning more money) or because one feels external pressure to do so (obligation to other, or as a mandatory subject in school) (Gardner, Tremblay, & Masgoret, 1997). It is due to the popularity of these concepts that the reasons an individual has for wanting to learn a language are commonly called *language learning orientations* (LLOs), rather than *goals*, which the term predominantly used in the Psychology literature (Dörnyei, 2003).

While the SE Model continues to be used in SLA research to this day, much criticism has also been levied against this theory, and particularly the integrative and instrumental orientation (Dörnyei, 2003; Leverett, 2012). Clément and Kruidenier (1983), for example, argued that the definitions of these LLOs are unclear and have thus been applied inconsistently in research. They noted, for example, that wanting to ‘travel abroad’, ‘having friends who speak English’, or ‘knowing more about English art, literature, and culture’ could be classified as either integrative or instrumental, depending on the researcher’s interpretation (p. 275).

Many other researchers have also pointed out that the integrative orientation in particular is problematic, because it is so highly dependent on the target language community. For example, where individuals are learning a language that is not spoken around them (i.e. *foreign language learning*), the desire to integrate with the target language community is likely to be less important than in the bilingual Canadian context in which the SE Model was developed. The English language in particular has become increasingly separated from its native speakers and cultures with the rise of globalisation and the concomitant spread of English as a lingua franca (Chen, Warden, & Cheng, 2005; Lamb, 2004).

In light of such criticism, a number of other language learning orientations were later proposed. Most notably, Clément and Kruidenier (1983) identified the *knowledge orientation* (a desire to learn a second language in order to gain broader access to information, pursue personal interests, and become better educated); the *travel orientation* (wanting to know a second language in order to be able to travel abroad); and the *friendship orientation* (a desire to foster interpersonal connections with people from other cultural or language backgrounds). The latter shows considerable conceptual overlap with another LLO that was suggested more recently by Yashima (2002, 2009): the *international posture*, which encompasses an ‘interest in foreign or international affairs, [...] readiness to interact with intercultural partners [...] and a non-ethnocentric attitude toward different cultures’ (Yashima, 2002, p. 57). Although this factor was first proposed to describe the motivation of Japanese learners of English as foreign language (EFL), its relevance to other contexts has been demonstrated extensively in research (e.g. Geoghegan, 2018; Kormos et al., 2011; Thurston, 2015).

Altogether, the introduction of the instrumental and integrative orientation can be seen as the most important contribution of the SE Model to the discourse on motivation in SLA. Together with the other LLOs which have been added since, they offer a straightforward way of studying the reasons have for language learning. Whereas it seems unlikely that instrumental goals would be strong motivators for engaging in informal second language practices, the integrative orientation, despite recent criticism, and its more recent counterparts the international and friendship orientation, seem more directly relevant to the ISLL context. For example, if target language communities are re-conceptualised as including non-native language users in addition to native speakers, identification with such communities may indeed be a more relevant motivator than ever before, as new technologies afford foreign language learners unprecedented opportunities for making connections with other target language users (Cole, 2015; Lam, 2006). In addition, the knowledge orientation also seems particularly pertinent to ISLL, given the many opportunities that informal second language practices offer for pursuing personal interests and seeking out information (see Section 2.1).

Furthermore, the SE Model suggests that individuals' attitudes towards the learning context determine the effort they are willing to invest in language-related behaviours. It seems that individual engagement in ISLPs may also, to some extent, be affected by attitudes towards this mode of learning (cf. Cole, 2015, Section 2.3.2). However, as attitudes were defined by Gardner with specific reference to the classroom (encompassing learners' opinions of their teachers and the methods used in their language course; Gardner, 1985b) it seems that this part of the SE Model would need some reconceptualization to make it applicable to the informal learning context.

Self-Determination Theory

While Gardner's SE Model arose from research within the field of second language acquisition, theories have also been brought into SLA from other areas of research. A leading example of this is the *Self-Determination Theory* (SDT), which was developed by social psychologists Deci and Ryan (1985). The application of this theory to the study of language learning motivation was spearheaded by Noels and colleagues (Noels, 2001a, 2001b; Noels, Pelletier, Clément, & Vallerand, 2000).

The most critical, and most widely known, part of the SDT is the distinction between *intrinsic* and *extrinsic* motivation. Intrinsically motivated individuals engage in certain behaviours due to the inherent enjoyment or satisfaction they gain from them, whereas extrinsic motivation implies that there are other reasons, which are external to the learner. Deci and Ryan (1985) maintained that 'self-determined' behaviours are more effective and sustainable, given that they are not dependent on the continued relevance of an external pressures or rewards. Contrary to popular belief, however, this does not mean that SDT espouses the superiority of intrinsic motivation over extrinsic motivation. This is because some forms of extrinsic motivation are considered to be as self-determined as intrinsic motivation, if they have been integrated thoroughly into an individual's sense of self (a process called *internalisation*).

Ryan and Deci (2000) identified several categories of extrinsic motivation along the continuum of internalisation. At the least internalised end of the spectrum (*external regulation*), an individual may be driven by instrumental benefits (e.g. obtaining good grades or a better job) or punishments (e.g. failing a class). Next, *introjected regulation* refers to cases where individuals have internalised external pressures to the point of feeling socially obliged to behave in certain ways (in order

to comply with others' expectations or fulfil perceived responsibilities). Even more internalised is the *identified regulation*, wherein a person is motivated to engage in behaviours which they feel are instrumental to achieving goals or performing tasks that are personally valuable to them (e.g. someone may decide to learn a few phrases in their neighbour's native language in order to better understand and connect with them). Finally, the most internalised form of extrinsic motivation specified in SDT is the *integrated regulation*, wherein individuals are motivated to pursue behaviours which they consider to be fundamental to their own identity or which enable them to be or become the kind of person they aspire to be (e.g. a young person in Europe may view learning and using a second language as an integral part of their identity as a European). Importantly, SDT also asserts that externally motivated behaviours may become more internalised over time (R. M. Ryan & Deci, 2000), if an individual recognises their personal value and/or incorporates them into their personal identity (e.g. develops a strong L2 user identity). That is, SDT introduced a more dynamic, situated perspective to the study of motivation, which until then had been understood as a fairly stable individual characteristic (Sugita McEown et al., 2014).

In terms of language acquisition, SDT focuses on the extent to which reasons for language learning are externally imposed versus integrated into the learner's identity and value systems. It can be used to gain insight into the perceptions that individuals who engage in second language learning have of themselves as language learners and users. Prior research has provided evidence for a link between ISLPs and identified and integrated motivation, that is, to learners perceiving their L2 as essential to carrying out activities they valued or reaching goals linked to their personal identity (Cole, 2015, see Section 2.3.2).

However, in SDT, self-perceptions are also inextricably linked learners' perceptions of the language itself and why it would be useful to know (i.e. learning goals). For this reason, initial research on SDT in SLA focused on mapping the four language learning orientations proposed by Clément and Kruidenier (1983) onto the intrinsic/extrinsic continuum Noels (2001b). It was concluded that there was considerable overlap between the two frameworks, with the knowledge and friendship orientation being, in general, more internalised than the more extrinsic instrumental orientation. Yet, the same LLOS do always map onto the same types of motivation in SDT: Two learners might at first sight have similar reasons for wanting to know another language but differ in the extent to which they have internalised these reasons. For example, one learner with a strong knowledge orientation may wish to learn an L2 because they feel limited in the information they can access in their L1 (*identified regulation*, according to SDT), whereas another learner may have a strong knowledge orientation because being a competent L2 speaker is central to their self-concept as a well-educated and well-informed individual (an example *integrated regulation*). Thus, SDT does not provide a clear-cut framework for studying learners' self-perceptions as separate from their perceptions of the language itself. However, as suggested by Noels (2001b) the notion of *self-determination* may be drawn upon to explain why certain reasons for language learning might be better predictors of effort and success than others.

The Second Language Motivational Self-System

Self-Determination Theory and other models which were adopted from the wider field of Psychology around the same time (see reviews by Al-Hoorie, 2017; Dörnyei, 2003) contributed to shifting the discourse on motivation in SLA away from the

primarily social and contextual factors emphasised in Gardner's (1985a) SE Model. Instead, researchers increasingly turned their attention towards learner internal-processes, particularly related to the self and identity, as well as understanding the situated and dynamic nature of motivation (i.e. how it is affected by the immediate learning context and changes over time) (Dörnyei & Ryan, 2015; MacIntyre, Mackinnon, & Clément, 2009a). These trends have given rise to other theories of language learning motivation (including, for example, Dörnyei and Ottó's [1998] *Process Model* of motivation and Ushioda's *Person-in-Context Relational View* [2009]) and culminated, most prominently, in Dörnyei's (2005) theory of the *Second Language Motivational Self-System* (L2MSS).

The L2MSS adapts the *possible selves theory* (Markus & Nurius, 1986) from Personality Psychology and applies it to the SLA context. This theory posits that the actions in which a person engages are, at least in part, motivated by future self-images, that is, how they imagine they might be or what they are afraid of becoming. Dörnyei's (2005) L2MSS thus includes two constructs which represent the language learners' imagined future identities: The *ideal L2 self* and the *ought-to L2 self*. The ideal self encompasses learners' aspirations, hopes, and wishes; 'the kind of L2 user that [they] would ideally like to be in the future' (Dörnyei & Ryan, 2015, p. 87). The ought-to self 'reflects attributes that one believes one ought to possess' (ibid.), for example to comply with others' expectations, fulfil perceived responsibilities, or avoid possible negative outcomes. In addition to the two future selves, the L2MSS also incorporates a third component, the *L2 learning experience* is intended to account for the impact of the specific learning context (e.g. the teacher and peers, the approach to learning and teaching, experiences of success or failure) on language learning motivation.

In the distinction between the fully internalised goals and aspirations which constitute the ideal L2 self and those of the ought-to self, which are perceived as having external origins, the L2MSS shows some theoretical overlap with the concepts of extrinsic and intrinsic motivation, although these are treated as a dichotomy rather than a continuum as in SDT (see above). Dörnyei (2005, 2009b) also suggested that the ideal self may be seen as an extension of Gardner's (1985a) concept of integrativeness, representing an individual's desire to become part of an imagined community of successful L2 users, which may or may not correspond to specific real-life communities of L2 speakers. The ought-to self, on the other hand, has sometimes been found to conceptually overlap with instrumental orientation (Noels, 2001a).

In the decade since its inception, Dörnyei's framework has been used in a plethora of research on second language learning (e.g. see reviews in Boo et al., 2015; Dörnyei & Ryan, 2015), most of which have focused on the more generalised concepts of the ideal and ought-to L2 selves and disregarded the L2 learning experience. Research results have consistently supported the L2 selves as valid and independent constructs in various (formal) learning contexts and have provided evidence for positive links with L2 learning behaviour and outcomes (e.g. Csizér & Kormos, 2009; Kim & Kim, 2014; Kormos et al., 2011; S. Ryan, 2009). It has been suggested that the L2 selves may only affect behaviour under specific conditions (Dörnyei, 2015; Oyserman & James, 2009), namely if they are *desirable* to the learner (they do indeed have future aspirations and goals), *accessible* (they are able to imagine their future selves), and *plausible* (the future selves seem achievable). However, these aspects have generally not been considered in research (Hessel, 2015).

The L2MSS has two characteristics which limit its applicability to informal learning contexts: First, the L2 learning experience, like Gardner's (1985a, 1985b) *attitudes towards the learning situation*, is defined with specific reference to features of classroom learning, operationalising the impact of the immediate learning environment on motivation via questions about 'the L2 teacher, curriculum, peer group, and the experience of success' (Dörnyei & Ryan, 2015, p. 88). The second major limitation of the L2MSS in the context of ISLL is the importance it assigns to *future-oriented*, as opposed to *current* self-perceptions. As previously noted by Cole (2015), individuals who are already successfully using their L2 to engage in informal second language practices may not consider the ideal future self very different from their current identity.

The last point is linked to one of the main criticisms that has been levied against the L2MSS (e.g. Hessel, 2015; Mercer, 2011): it is lacking a component which represents the learners' current L2 self-concept, despite proposing (following Higgin's [1987] *Self-Discrepancy Theory*) that the distance between the current and future selves is key in motivation. MacIntyre, Mackinnon, and Clément (2009b), however, argued that learners who are already successful language users (individuals with both a strong current and future L2 self, i.e. low discrepancy at the higher end of the scale) could be equally motivated to engage in language-related behaviours, pursuing 'ongoing development' rather than a far-off, future vision of themselves. In light of this theory, and with the addition of a *current L2 self* measure, the L2MSS seems to have the potential to provide insight into the role of learner self-perceptions in ISLL.

Despite the recent popularity of the Second Language Motivational Self-System (Boo et al., 2015), there is still not one predominant theory of motivation in SLA. As noted by Sugita McEown and colleagues (2014), different elements from these theories are useful in explaining different aspects of behaviour and predicting learning outcomes (see also Csizér & Kormos, 2014; Iwaniec, 2014; Kormos et al., 2011). As stated in the introduction to this section, the goal in incorporating measures of motivation into the current study was to gain an understanding the relationships between engagement in informal second language practices and individual learners' perceptions of their second language and why it would be useful to know (goals), of the learning context (attitudes), of the individuals themselves as language learners and users (self-concept). The rest of this section discuss how each of these aspects of motivation could be operationalised in the current study using the theories discussed above.

Goals

All three of the theoretical frameworks introduced above include components which evaluate language learners' *goals*, that is, their perceptions of the language itself and why it would be useful for them to know. However, both the SDT and L2MSS consider reasons for language learning exclusively through the lens of how strongly they have been integrated into the learners' identity, therefore conflating goals and self-concepts. By contrast, the language learning orientations that emerged from the SE Model (Clément & Kruidenier, 1983; Gardner, 1985a; Yashima, 2009), seem to offer a way to separately operationalise learners' reasons for language learning in the current study.

Attitudes

Attitudes towards the learning context are included as separate components both in the SE Model and the L2MSS. However, both frameworks aim to motivation specifically in formal language learning context, and thus operationalise attitudes with specific reference to features of classroom settings, including learners' opinions of their teachers and peers, the curriculum, and the methods used in their language courses (Dörnyei, 2005; Gardner, 1985a, 1985b). While it has been suggested that attitudes towards formal learning may also have an impact on learners' behaviours beyond the classroom (e.g. Chik, 2014, 2018; Cole, 2015; Godwin-Jones, 2019, see Section 2.3.2), the researcher decided that a measure of learners' attitudes towards *informal* learning would also need to be included in the current study.

Conceptions of the self and identity

Finally, Dörnyei's (2005) theory of the L2 selves seems, with some changes and additions (see above) to be the most suited to operationalising learners' *self-perceptions*. In particular, the ideal L2 self adds an interesting temporal dimension by considering the extent to which learners imagine themselves as a successful user of their target language in the future. The researcher thus decided to include this variable in the current study with the addition of a complementary *current L2 self*, measuring of the extent to which the participants already perceived themselves to be competent users of their target language (Hessel, 2015; Thurston, 2015). The current L2 self bears some resemblance to other concepts used in research on motivation in SLA, such as *self-confidence* and *self-efficacy* (Dörnyei, 1994), and allows for a comparison of the relative roles of current user and learner (i.e. future user) identities in ISLL in the current study.

The construction of the instrument used to measure motivation and attitudes in this study will be discussed in Section 6.1.2. Appendix B.2.1 contains an overview of the constructs included in this study, their definitions, and the questionnaire items through which they were operationalised.

The following section contains an overview of previous studies which have investigated motivation in relation to informal second language learning.

2.3.2 Motivation and Attitudes in Informal Second Language Learning

The most extensive study of motivation in ISLL to date was conducted by Cole (2015, for a summary of the design, see Section 2.1.2). In addition to asking about their attitudes towards formal and informal learning, Cole used the Self-Determination Theory taxonomy of motivation (see Section 2.3.1) to code the participants' present and past motivation for learning English. The results revealed notable differences in the motivational development of the 'classroom-trained learners' (CTLs) and 'fully autonomous self-instructed learners' (FASILs).

Most of the participants in both groups reported that, at the beginning of their learning trajectories (past motivation), they did not have the explicit goal or desire to learn English (they were *amotivated*). Nevertheless, they continued to learn English after they finished the compulsory foundational classes at school. Most of the CTLs, who at the time of Cole's (2015) study were enrolled in language classes at prestigious private language schools, did so because they felt pressured by their parents (external regulation). By contrast, FASILs continued to learn English mostly incidentally, as a by-product of engaging with informal resources which related to their personal interests (identified regulation).

Meanwhile, at the time of Cole's (2015) study (present motivation), almost half of the CTLs (43%) were still showing an external regulation, although 30% said they were now more intrinsically motivated (i.e. by a general fondness for English and enjoyment of learning). Meanwhile, 75% of the FASILs reported some form of identified or integrated regulation (which Cole collapsed into one category, 'internal extrinsic motivation'); that is, they perceived English as essential to carrying out activities they valued (e.g. reading English-language books or watching TV series and films) or reaching goals linked to their personal identity (becoming part of a global community and/or developing an international identity).

These findings suggest a relationship between motivation and informal activities, since the FASILs spent significantly more time on ISLPs than the CTLs (no direct correlations for motivation and ISLPs were reported by Cole). Furthermore, whether or not participants displayed 'internal extrinsic motivation', together with FASIL group membership, was the primary predictor of English language proficiency in this study. Cole interpreted this as evidence that these learners were 'acquir[ing] compelling user identities based on belonging to authentic communities of practice and using the language for authentic reasons' (2015, p. 334).

Lyrigkou (2016) also used Self-Determination Theory to study Greek secondary school students' general motivation for learning English and their motivation for engaging with English through informal activities. Her participants' patterns of general language learning motivation were somewhat similar to those among Cole's (2015) classroom-trained learners: Approximately equal numbers of her participants were extrinsically motivated to learn English (43%) as displayed an identified or integrated motivation (49%). Somewhat surprisingly, external reasons were even

more dominant in her participants' motivation to engage in ISLPs: 62% of the surveyed pupils reported engaging in informal activities primarily to improve their language skills to do better on exams; 20% viewed ISLPs as essential to carrying out activities they valued (integrated-introjected regulation); and 16% were motivated intrinsically by their enjoyment of learning and using English.

Finally, Lyrigkou (2016) examined which category of motivation to engage in ISLPs was most strongly associated with actual informal contact. She found a small, significant correlation between the frequency of informal contact and identified motivation ($r_s = .30, p < .01$), supporting the (perhaps self-evident) conclusion that students who were motivated to engage with English in their free time because it allowed them to carry out personally valued activities (e.g. playing video games, reading in English) spent more time engaging in ISLPs than students who did not see these activities as central to their personal interests.

Instead of focusing on general language learning motivation, Lai, Hu and Lyu (2018) explored the relationships between students' perceptions of the usefulness and ease of use of technology-assisted out-of-school activities and the extent to which they actually engaged in them. The researchers used Structural Equation Modelling to analyse self-report questionnaire data from 439 learners of various L2s in Hong Kong and the USA. The study separately investigated three types out-of-class second language practices: self-instruction (i.e. non-formal) activities, social activities, and entertainment- and information-oriented practices. Most relevant to the current review are their findings pertaining to the last category, which encapsulate the same types of activities as *informal second language practices* in the current study.

The frequency with which the participants engaged in entertainment- and information-oriented activities was predicted equally by their beliefs about the usefulness of these activities for language learning and their perceived ease of use (both $\beta = .19, p < .10$). These perceptions were, in turn, significantly affected by teacher and peer attitudes. Lai, Hu and Lyu (2018) also tested their models separately for beginner and advanced learners (self-reported proficiency below intermediate level vs. above) and found that perceptions of the usefulness and ease of use of informal activities only significantly affected the practices of learners at higher proficiency levels (respectively: $\beta = .23$ and $.20, p < .10$). This suggests that informal second language practices may serve different functions at different stages of language acquisition, and that different approaches may need to be developed to encourage engagement in such activities among beginner and more advanced learners.

Whereas Cole (2015), Lyrigkou (2016), and Lai, Hu and Lyu (2018) tried to understand what drives students to engage in informal second language learning, Lamb and Arisandy (2019) set out to explore the impact of frequent informal language contact on motivation, particularly the motivational L2 selves and attitudes towards classroom learning.⁶ In terms of attitudes, Lamb and Arisandy (2019) hypothesised that frequently engaging with English outside of school might lead students to perceive an ‘authenticity gap’ (as suggested by Henry, 2013), that is, to value these activities more highly than their formal language instruction. In terms of

⁶ It should be noted that all of the studies except for Lai, Hy and Lyu’s (2018) research employed correlational methods, so that no conclusions can be drawn on the basis of these findings about the directionality of the relationship between ISLPs and motivation.

language learning motivation, however, the researchers hypothesised that informal language contact would have a *positive* effect on students' motivation to learn English.

Using self-report questionnaires, Lamb and Arisandy (2019) collected data from 308 university-level EFL learners in Indonesia about their language learning motivation and the English online activities in which they engaged in their free time. Following Lai, Hu & Lyu (2018, see above) these were categorised as either self-instruction, social or entertainment-oriented. Again, the results pertaining to the latter category are the most relevant to this review, as it roughly encompassed what are called ISLPs in the current study.

Whereas the results indicated no significant links between any form of out-of-class language contact and the ought-to L2 self, the ideal L2 self correlated significantly with all three categories of informal practices (entertainment-oriented: Pearson's $r = .27$; social: $r = .29$; self-instructed: $r = .39$, all $p < .01$). This supports the researchers' hypothesis that ISLPs can increase students' internalised motivation for English learning (although it may equally be interpreted as showing that students with a stronger ideal L2 self are more likely to engage in ISLPs). Meanwhile, there were small but significant *positive* correlations between social and self-instruction oriented practices and attitudes towards formal learning (respectively: $r = .15$ and $r = .21, p < .01$), and *no* significant relationship between attitudes and entertainment-oriented activities. That is, the data did not support the hypothesis that frequent engagement in informal second language practices negatively affects learners' attitudes towards formal instruction. In fact, the positive correlations suggest that learners engage more frequently in L2 activities outside of the classroom hold slightly more positive attitudes towards formal learning as well, or conversely

that positive attitudes towards formal learning may affect students' engagement with their second language even beyond the classroom.

In fact, a number of prior studies have found evidence which suggests that attitudes towards formal language learning has may have an effect on informal practices beyond the classroom (Godwin-Jones, 2019), although not in terms of the quantity of ISLPs, but the types of interactions students have with them (e.g. cognitive and linguistic engagement). For example, Chik (2014) observed that university students in Hong Kong, who frequently played video games in English, intentionally applied strategies they had been taught in school to turn their gaming experiences into opportunities for L2 learning. Chik (2018) later confirmed this finding in another sample of adult learners of German, Italian, and Korean in Hong Kong, who she found were deliberately choosing resources and employing language learning strategies to try to re-create more structured classroom learning experiences.

Evidence of a slightly different kind was provided by Lai (2015), who also studied the practices of L2 learners in Hong Kong and found that students with different learning beliefs used the same non-formal resources in different ways. For example, she observed that students who believed in the necessity of systematic grammar learning often used grammar websites as self-study tools, whereas learners who did not believe in studying grammar explicitly would only sometimes consult such sites to help them resolve comprehension issues, for example while watching TV series in their L2.

Finally, a small number of studies have explored the relationships between informal second language practices and other affective variables which bear some resemblance to concepts used in the current study: learners' enjoyment of language learning, which

is part of intrinsic motivation, and their confidence in learning English, which is similar to the current L2 self in that it measures learners' perceptions of their own linguistic capabilities, but differs in that it emphasises their status as *learners* rather than *users* of the language.

As reported previously (Section 2.1.1), both Lai et al. (2015) and Lee (2019) found medium-sized positive correlations between students' confidence in and enjoyment of learning English and the *variety* of their out-of-class L2 activities. In the former study (Lai et al., 2015), the variety of meaning-focused activities, together with the frequency of participants' use of technology, explained 20% of the variation in their confidence and 30% in enjoyment. However, In Lee's (2019) research, the variety of ISLPs was no longer a significant predictor of confidence or enjoyment after demographic variables (age and major) were also included as co-predictors in the same multiple regression model.

The amount of time students spent on informal second language practices was not significantly correlated to confidence or enjoyment in the research by Lai and associates (2015). In Lee's (2019) study, on the other hand, there were significant correlations of quantity with both variables and, unlike the variety of activities, the time participants spent on ISLPs remained a significant predictor of confidence and enjoyment even after the inclusion of the demographic variables. Together, the participants' age, major, and the quantity of ISLPs predicted 23% of the variance in confidence and 27% of the variance in enjoyment.

The wide variation in the findings of the studies reported above suggests that the relationship between motivation and attitudes and informal second language practices is complex. More research is needed, for example to explore whether

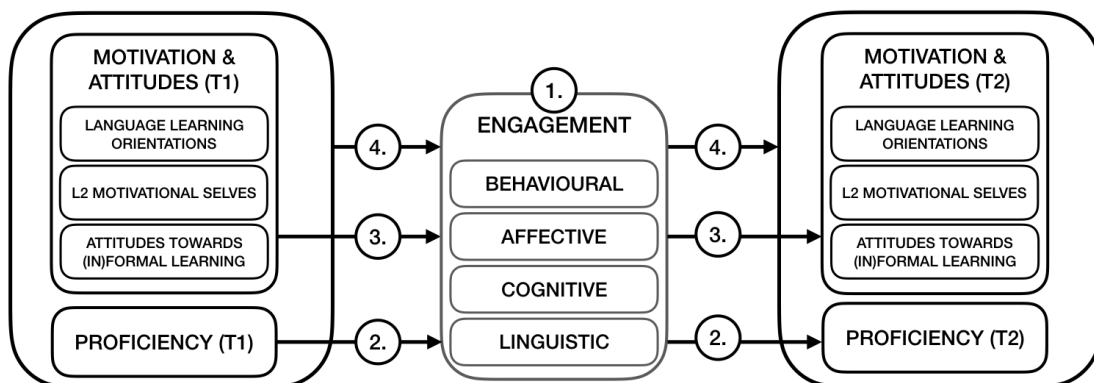
particular informal activities have a greater potential to promote confidence, enjoyment, and motivation in language learners than others, and under which conditions, or indeed whether more enthusiastic and confident learners are more likely to engage in a greater variety and quantity of informal second language practices. Thus, one of the aims of the current study was to gain a better understanding of the role of motivation and attitudes in ISLL.

3. Methodology: Introduction to the Study

3.1 Theoretical Model and Research Questions

This study was designed based on the hypothesis that the ways in which individuals *engage* in informal second language practices might help to explain some of the variation that has previously been found in language development and changes in motivation associated with informal second language learning. The theoretical model (Figure 3.1) thus placed engagement as a multidimensional mediator between language proficiency and a number of motivational and attitudinal variables, measured at two time points.

Figure 3.1: Theoretical Model — Motivation, attitudes, proficiency and engagement in informal second language learning.



The main research questions (RQs) were designed to explore, step-by-step, the relationships indicated in the model above (numbered 1–4):

1. How do learners of English as a second language, attending year 10 classes at German *Gymnasien*, engage in English informal second language practices?
2. What are the direct relationships between language proficiency and engagement in informal second language practices?

3. What are the direct relationships between language learning motivation and attitudes towards language learning and engagement in informal second language practices?
4. How do language proficiency, language learning motivation, and attitudes interact in their relationships with engagement in informal second language practices?

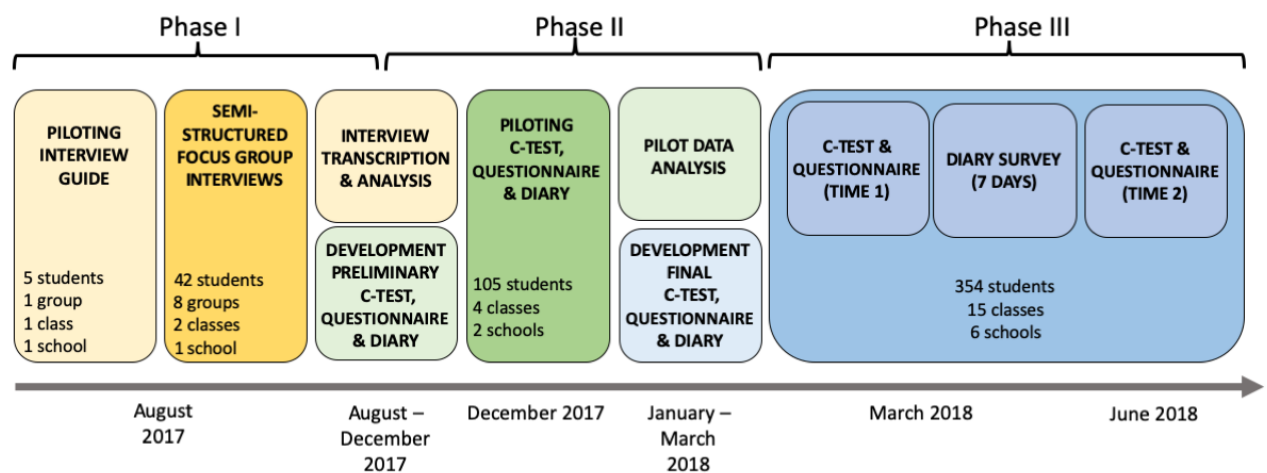
For the purposes of this study, engagement was defined as ‘the observable and unobservable qualities of students’ interactions with learning activities’ (Wang et al., 2016, p. 17). That is, it served as a way to operationalise the nature (or *qualities*) of informal L2 practices in terms of four dimensions: behavioural, affective, cognitive, and linguistic engagement (see Section 2.2). As this study was not focused on determining the specific types of linguistic and communicative competencies which learners may acquire through ISLL, changes in language proficiency were evaluated globally through C-tests (see Section 6.1.1). The motivational and attitudinal variables which were investigated included language learning orientations, L2 motivational selves, and beliefs about informal and formal learning (see Section 2.3).

3.2 Research Design

To answer the research questions stated above, a mixed methods study was designed, which consisted of three data collection phases. Phase I (Chapters 4 and 5) involved the collection of qualitative data through semi-structured focus group interviews. During these interviews, students from the target population were questioned about how and why they engage in ISLPs and about their attitudes towards informal learning. The findings from these interviews informed the subsequent development

of questionnaires to measure motivation, attitudes, and engagement in ISLPs in Phase II (Chapters 6 and 7). At this point, a C-test was also designed specifically for use with this target population. All instruments were piloted with members of the target population and refined based on the results from the pilot data analysis, before moving on to Phase III (Chapters 8 and 9). This last stage of the project involved the collection of quantitative data from another, larger sample of students. The participants completed C-tests and motivation and attitudes questionnaires at two time points, three months apart. They were also asked to use an online survey to record their informal language contact for one week immediately following the first data collection. Figure 3.2 provides an overview of the three-phase research design and timeline.

Figure 3.2: Research design and timeline



The rationale for these design choices will be discussed further in the methods chapters for each phase. However, there were a number of considerations in the overall research design which bear mentioning at this point.

Questionnaire and proficiency data were collected at two time points, immediately before the learner diary and 14 weeks thereafter, to allow for the evaluation of changes over time. That is, the model at the heart of this study (Figure 3.1) focuses on investigating both the extent to which prior language proficiency, motivations and attitudes can predict subsequent engagement in ISLPs, as well as whether engagement can predict later language proficiency scores (while controlling for proficiency at time 1 [T1]). It is important to state at this point that, the model tested in this study assumes directionality of the relationships between variables based on their temporal order, it ultimately cannot settle the debate about the directionality of the relationship between language proficiency and informal second language practices. To truly capture the ‘complex dynamic relationship’ between these variables, data would have to be collected more frequently, over a longer period of time (cf. De Bot & Larsen-Freeman, 2011). However, because of the high burden on participants, such studies are feasible only with significantly smaller samples (Kusyk, 2017a).

There were two main reasons for choosing to collect data across a 14 week interval in this study: First, the researcher wanted to evaluate language development related to ISLPs over longer periods of time than in previous studies (Kusyk, 2017a; Sundqvist, 2009, 2011), in which repeated measures were collected only 6–8 weeks apart. Secondly, in consulting with the schools that were interested in participating, it was determined that the most convenient time for the data collection to take place would be immediately before the start of the Easter and summer vacations (mid March and late June, i.e. 3 months apart). This served to minimise the interference of the study with the participants’ regular curriculum learning, exam preparations, and

test periods and allowed the researcher to recruit the large number of participants needed for this study (see below).

It decided that the relationships between engagement, general language proficiency, and the various motivation and attitudes variables over time would best be explored systematically explored through structural equation modelling (SEM). This is because SEM allows for simultaneous testing of complex inter-variable relationships, including mediational models such as represented in Figure 3.1. Second, SEMs can be used to adjust for measurement where, instead of directly measuring a target construct, instruments such as multi-item scales are used as indirect indicators of latent variables, as was the case with almost all of the constructs in this study (including proficiency, motivation, and some aspects of engagement).

Since large data sets are needed for SEM, the researcher strove to minimise the burden on the schools and participants in order to facilitate recruitment. By choosing time-efficient C-tests (Section 6.1.1) and self-report questionnaires (Section 6.1.2) as instruments, data on proficiency, motivation, and attitudes could be collected within one school lesson (45 minutes). Information about engagement in ISLPs, on the other hand, was collected using an event-contingent informal language learning diary (Section 6.1.3), which the participants were asked to keep for seven days. The researcher recognised that this placed a high demand on the students in terms of time and self-directed effort at home. However, this method was ultimately chosen (over one-off, retrospective self-report questionnaires, which have often been used in previous ISLL research), as it was thought that it would provide a more accurate record of learners' experiences and especially their concurrent cognitive and emotional states.

Of course, the degree of generalisation inherent in large-scale statistical modelling such as SEM may seem to stand in contrast to the overarching goal of this research, which was — through the choice of the central constructs engagement, motivation, and attitudes — to focus on understanding the individual differences between learners involved in ISLPs. With this in mind, two higher-level design choices were made specifically to reflect the theoretical focus on the individual learners: Firstly, it was decided to take a bottom-up approach to developing the quantitative research instruments, that is, to start by speaking to individuals from the target population. The goal of the focus group interviews (Phase I) was to explore what the learners themselves thought were the most relevant aspects of their engagement in ISLPs, their motivation for engaging in these practices, and attitudes towards this mode of learning. This was then taken into consideration when designing the research instruments (Phase II) and used for triangulation when interpreting the quantitative findings (Phase III).

Secondly, a more learner-focused perspective was introduced in Phase III by choosing Latent Profile Analysis (LPA) as one method for analysing the engagement data. LPA is a type of statistical modelling, which seeks to identify distinct groups (or *profiles*) among a sample of participants based on covariance between a number of observed variables — in this case the different dimensions of engagement. As such, LPA is inherently more learner- than variable-focused (as most other statistical tools are, including SEM). It was hoped that exploring different engagement profiles among the participants in this research would contribute to understanding what individuals are actually doing, thinking, and feeling when they participate in ISLPs; the extent to which this varies between different (groups of) learners; and how it relates to proficiency, motivation, and attitudes.

3.3 Target Population and Recruitment

The target population of this study were pupils in year 10 English language classes in Germany. There were four reasons for choosing Germany as the setting for this research:

1. Because students there are likely to be technologically well connected, and thus have plenty of opportunities to engage in ISLPs.
2. Because television and films in Germany are dubbed, not captioned, so that English is bound to be less naturally prevalent in these learners' lives than in most of the previous ISLL studies, which were conducted in countries which predominantly show English-language films and television programmes with the original soundtrack and translated subtitles.⁷
3. To expand the scope of ISLL research, while maintaining some ability to compare results: Previous studies have been conducted with secondary school students in other countries (e.g. the Netherlands [Kuppens, 2010] and Sweden [Sundqvist, 2009]), as well as with university students in Germany (Kusyk, 2017a).
4. For reasons of access: The research is from Germany and her first language is German.

To control for variation in formal language education — inasmuch as is possible in a large-scale study such as this — the participants were only recruited from one type of secondary school (*Gymnasien* – roughly equivalent to British grammar schools). *Gymnasien* were chosen because pupils in this school form — the only one to prepare

⁷ E.g. Sweden (e.g. Olsson, 2011; Sundqvist, 2009, 2011; Sundqvist & Sylvén, 2014), Belgium (Kuppens, 2007, 2010), China (Lai et al., 2015), and the Netherlands (M. Verspoor et al., 2011).

its students for entry into higher education — might be more likely to show a wider range of engagement in informal English language activities, given that they are likely to be more proficient and that the subject English can be considered highly relevant to their academic studies and future success. Year 10 students were recruited for this study for several reasons: Firstly, 15- to 16-year-old pupils are likely have a high degree of autonomy in their recreational activities and their internet usage. Secondly, year 10 is the final year in which all students at *Gymnasien* are taught English together, before being setted into higher- and lower-level courses to prepare for end-of-school exams.

In addition to restricting recruitment to one type of secondary school, it was also limited to schools located in only one of the German federal states, Lower Saxony (which consistently scores averagely on the government's annual assessment of educational outcomes [“Wo stehen die Bundesländer beim Bildungscheck?” 2016]). Since subject curricula in Germany vary by state and type of school (König, 2018), limiting recruitment to *Gymnasien* in Lower Saxony means that all participants had been taught to the same curriculum, although there were still likely to be some variations in teaching styles between different classes, schools, and teachers.

To somewhat control for the students’ socioeconomic status (which is a known predictor of academic outcomes [e.g. Ariani & Ghafournia, 2016; Filmer & Pritchett, 1999; Sirin, 2005] and is thought to influence access to informal resources [Sundqvist, 2009]), the participants were all recruited from rural and provincial areas, specifically from administrative districts surrounding the city of Bremen (national data show that there is less socioeconomic disparity within the area surrounding large cities in Germany, including Bremen [Albrech, Fink, & Tiemann, 2015]). In addition to these sampling restrictions, data exclusions were used as another method of controlling for

important factors that may influence informal second language learning. Specifically, to control for L2 experience, participants who were enrolled in bilingual (English-German) programmes, who considered English their native language, or who had spent more than a total of six weeks living in a primarily English-speaking environment were excluded from the analysis.

The study employed a volunteer sampling method: Suitable schools in the target region were identified and approached via email, phone calls, and through personal contacts with the aim of recruiting whole classes. They were supplied with information (in German) about the study, including the aim and details about what would be required of participating schools, teachers, and pupils. First, contact was made with the principals and later, with their permission, with year-10 English language and/or class teachers. Those who agreed to take part in the study distributed information letters to their pupils (again, in German), accompanied by a verbal explanation of the project and an opportunity for students to ask questions about the project. Written consent was obtained from both students and parents; the collection of reply slips was managed by the teachers. If a sufficient number of students from one class declared their interest in taking part in the study, a suitable time for the data collection was arranged with the teachers.

The samples of pupils who participated in each phase of the research were partially overlapping (i.e. some classes participated both in Phase I and II, but there was no overlap between Phase III and any of the previous phases). The subsamples will be described in more detail in the respective methods chapters.

3.4 Ethical Considerations

This project was designed following the ethical guidelines proposed by the British Educational Research Association (“Ethical guidelines for educational research,” 2011). Prior to each phase of data collection, the methods were approved by the University of Oxford Central University Research Ethics Committee (CUREC) as well as the Lower Saxony State Education Department. Recruitment commenced after approval from both bodies was secured.

If a teacher agreed to take part in the study, all students in the class and their parents/guardians received comprehensive written information about the study. Signed consent forms were collected from those who decided to participate. The participants and parents/guardians had several days in which to decide whether or not they wished to participate, to avoid making them feel pressured (in view of the unequal power relations between the students and their class or subject teachers). They were also invited to contact the researcher with questions before deciding to participate. No rewards were offered for participation. It was made clear that this study was completely unconnected to the students’ academic assessment and progress, and that their data would not be disclosed to the teachers or the school. In addition to this written information, at the start of each phase of data collection, the pupils were given a brief oral introduction to the study and were reminded again of the anonymity and confidentiality of their responses and their right to withdraw at any time.

As mentioned previously, the interviews, tests, and survey were designed to take up as little class time as possible and steps were taken to minimise the interference of the study with regular curriculum teaching, exam preparations, and test periods. Where either pupils themselves or their guardians did not give their

consent for participation, the teacher prepared alternative learning activities for the students (e.g. self-study exercises) to make up for the lost class time.

Every effort was made to keep the participants' data anonymous and secure during all stages of the project. While travelling to and from research sites, the paper-based data (questionnaires and C-tests) and devices containing digital data were kept in a locked piece of luggage and closely monitored to prevent loss or theft. Once the researcher returned to the UK, these papers and devices were securely stored in a locked filing cabinet at the University of Oxford Department of Education. After converting the data into electronic format, the papers were shredded and the digital recording devices' memory permanently deleted. All digital data files were kept on the researcher's computer in an encrypted, password-protected folder and a digital backup (also encrypted and password-protected) was created through the university backup service.

In designing the language learning diary, which took the form of an online survey that participants completed outside of school, on their mobile phones or computers, the researcher followed the CUREC Best Practice Guidelines (*Internet-based research (IBR): Best practice guidance 06_Version 3.0*, 2016). Accordingly, the online survey did not collect participants' IP addresses, and the software used ("Qualtrics Research Suite," 2018) adhered to the EU-US Privacy Shield data protection framework, meaning that the processes for transferring of data from servers located in the European Economic Area to the US were compliant with European privacy laws (*Privacy Shield List*, 2018).

Although every reasonable effort was taken, complete safety during data transfer (from the participants' phones to the Qualtrics servers, and to the researcher's computer) could not be fully guaranteed. Therefore, the written information for

participants and guardians contained a disclaimer to this effect. As an additional precaution, the participants were instructed that they should use a secure, private internet connection to ensure their digital safety. To protect participants in the unlikely case that the internet connection or survey database was made unsafe, all data collected online was anonymous and not traceable to the individual pupil (i.e. no identifiable participant data were retained; every student was assigned an ID number to allow the collation of the diary entries with the remaining data).

Finally, thorough ethical considerations were involved in sending the participants reminders to their mobile phones in Phase III (Section 8.2). Consent to collect the participants' phone numbers was sought from the teachers, the students and their parents/guardians. The phone numbers were kept apart from any other, potentially identifying information: The students wrote their phone numbers on a separate sheet of paper, without their names, and they were saved on the researcher's phone under randomly-assigned identifiers (e.g. 'No. 1') that were different from the participant IDs used in the rest of the study. One reminder message was sent to each number daily for seven days, and then the phone numbers were immediately and permanently deleted. The reminders were sent separately to each individual student, rather than setting up a group chat, to keep the pupils' numbers private from each other. However, some classes, who already used a WhatsApp group to communicate with each other and with teachers, chose instead to identify one volunteer to receive the daily messages from the researcher, which were then forwarded to everyone via the group chat.

4. Methods Phase I: Qualitative Data Collection

The first phase of the study involved the collection of qualitative data through semi-structured focus group interviews. This was intended to help the researcher explore the types of informal practices popular among students from the target population, how they themselves conceived of these activities (e.g. as entertainment, opportunities for learning, etc.), and the beliefs they held regarding learning English as a second language in general and informal learning in particular.

4.1 Participants

The participants in this phase of the study (including piloting) were 47 pupils from one *Gymnasium* in the rural district of Oldenburg, Lower Saxony. To improve the transferability of the findings from this stage, the initial aim had been to interview participants from more than one school, but recruitment proved difficult due to the time-intensive nature of the interviews. However, the participants did belong to two different year 10 classes and were taught English by different teachers, which allowed the researcher to reflect on the influence of factors such as the class dynamics and teaching style.

All participants were between 15 and 16 years old ($M = 15.2, SD = .40$). Among them were 37 girls (78.7%) and 10 boys (21.3%). There are two main reasons for the gender imbalance in this sample. First, while there is generally a slight prevalence of girls in *Gymnasien* (in 2015/2016, 53.8% of the students at these secondary schools in Lower Saxony were girls [Die niedersächsischen allgemein bildenden Schulen in Zahlen, 2016]), the headmaster of this school also commented that this particular year group contained an ‘uncommonly high’ proportion of girls (exact numbers not obtained). Second, the researcher found that the boys in the

participating classes, more so than the girls, did not want to take part in the study because they were unsure what to expect (these pupils had never taken part in research before). This was evidenced by that fact that, in the end, four more boys decided to participate in the study (with parental consent) than had initially volunteered, because their classmates, who had already been interviewed, had spoken positively to them about their experiences.

Six participants indicated that they had grown up bilingually, speaking German and one other language. Only one of them was a German-English bilingual speaker. The remaining 41 participating pupils were German monolinguals. Of the 46 participants for whom English was not one of their native languages, the majority (37; 80.4%) had started learning English at age eight, or in their third year of primary school. This is when the language is most commonly introduced in Germany. The rest (8; 17.4%) had started to learn English at age six or seven (in their first year of primary school). That is, at the time of the study, most of the participants had been learning English formally for seven or eight years ($M = 7.64, SD = 1.21$). Their teachers estimated that, according to the Common European Framework of Reference (CEFR), their pupils had, on average, a B1-B2 level of competency (Council of Europe, 2001). This corresponds to what the curriculum for Gymnasien in Lower Saxony stipulates students should achieve in year 10 (Böwing, Frome, Gerking, Haupt, & Schulte, 2006).

4.2 Procedure

There were two reasons for choosing focus group interviews as the main qualitative data collection method: First, compared to one-on-one interviews, focus groups are an efficient way of gathering data from more participants. A greater number of

interviewees allows the research to capture a wider range of thoughts, feelings and experiences within the target population. Second, there is an assumption that the interaction between the participants, as well as being part of a group (not feeling 'put on the spot') will elicit richer data than individual interviews, particularly where the goal is to explore the diversity of thoughts, feelings, and perceptions among individuals (e.g. Ho, 2006; Krueger & Casey, 2015). Focus group interviews do not seek a consensus, but to facilitate discussion among the participants. Listening to the contributions of others may help the interviewees identify and reflect on their own experiences and opinions and bring to light ideas that may not emerge in one-on-one settings (*ibid.*).

Of course, the chosen data collection method was not without potential limitations. For example, many aspects of ISLL may be unconscious, in which case the participants might not have been able to accurately report their practices or articulate what they felt and thought about them. Additionally, much like in one-on-one interviews, there was a risk of the researcher's presence leading to bias in the focus group participants' responses: In this study, which involved a researcher from the University of Oxford (which is perceived as exceedingly prestigious in Germany) interviewing 15-year old secondary school students, the participants may, for example, have felt the need to 'intellectualise' their experiences, or well-meaningly give answers which they thought would appeal to the researcher or other participants in the group (*cf.* Krueger & Casey, 2015). In some cases, the presence of other students, and their reactions to what their classmates were saying, helped the researcher to recognise when this was happening. Finally, since the researcher acts as more of a moderator than interrogator during focus group interviews, the risk of

leading or biasing the interviewees is arguably reduced compared to on-on-one settings.

Nevertheless, the researcher attempted to minimise the potential impact of the unequal power dynamic in several ways: Firstly, a small common room in a quiet part of the school — open to use by students from year 10 and above for group work and after-school clubs — was chosen as the location for the focus group interviews because it was familiar to the participants. It was also thought that this room would provide a safer space for the pupils to talk about any aspects of their formal and informal learning experiences than, for example, a classroom, because it was not directly associated with the core academic aspects of their school life. Furthermore, during the interviews, the participants and the researcher were seated around an oval table, allowing everyone to look at everyone else while they were speaking, and for the researcher to be seated among the participants as their equal, not across from them in an interviewer position. Finally, the researcher consciously focused on building rapport with the students (e.g. by relating to them over specific pieces of popular culture, such as YouTube channels or TV series) and otherwise tried to ‘stay in the background’ as much as possible, taking on the role of moderator: asking questions, listening attentively, keeping the conversation on track, and making sure that everyone had a chance to share their opinions instead of letting particularly vocal participants dominate the conversation.

The focus groups were conducted as semi-structured interviews, that is the conversation was guided by questions that the researcher had designed a priori, but the order of the questions and wording was adjusted as needed. Sometimes, questions were also added ad hoc, according to how the conversation progressed, for example to prompt more detailed answers or to explore previously unexpected narratives. The

interview guide (Appendix A.1) was designed to elicit details about, and highlight the most relevant aspects of, the respondents' engagement in ISL practices (including behavioural, cognitive, affective, and linguistic engagement), their motivation for engaging in these practices and for language learning in general, as well as their attitudes towards informal and formal modes of learning. The questions were meant to be exploratory, that is, to offer the interviewees a chance to speak about aspects of their ISLPs of which the researcher may not have been aware. For this reason, the questions were open-ended, enabling the participants to share their experiences and opinions freely, without being limited by predetermined answers. However, for some questions (e.g. *'Can you think of occasions where you communicate by writing or speaking English?'*) the researcher offered the participants examples of possible answers as prompts when needed. The questions were grouped thematically, moving from general to more focused to help the natural flow of the conversation.

The interview guide was piloted with one group (five students), who were also asked to comment on their ease of understanding the questions and whether they felt that any questions were missing. Their answers to and feedback about the questions was used to finalise the interview guide and to determine whether the included questions would generate sufficient and relevant responses. After minor adjustments, mostly regarding the wording of certain questions, the finalised guide was used with the remaining eight focus groups.

The composition of the individual groups was decided by the pupils together with their teachers. The groups consisted of four to six students, a group size which is often recommended to strike a balance between being small enough to give every interviewee a chance to participate and large enough to include a variety of viewpoints to spark discussion (cf. Krueger & Casey, 2015). Wherever possible, the

focus groups were formed around pre-existing affinity groups, in the hope that being with their friends would make the participants feel more at ease and free to speak their mind. Since gender is one of the main individual variables in this research, interviewing boys and girls separately was considered at first, but eventually the researcher decided against it: The small number of boys among the participants made this division difficult and the students' preferred affinity groups were mixed gender.

Over the course of one week, the pupils took turns leaving their regular classes for 90 minutes (one double lesson) to participate in the focus groups. Before commencing the interview, the researcher gave a brief oral introduction to the study and answered any remaining questions. Depending on the group, the interviews lasted between 40 and 60 minutes. The discussions were audio-recorded using two portable digital recording devices and later transcribed by the researcher for analysis. After all questions from the interview guide had been answered and the conversation had reached a point of saturation, each participant also filled in a brief questionnaire (Appendix A.2) to provide some demographic data to help interpret the findings.

4.3 Analysis

Since only minor adjustments had to be made to the interview guide, and semi-structured interviewing allows for a certain degree of flexibility, the responses from the pilot interview were analysed alongside the data from the other eight focus groups. The transcripts were analysed using *directed content analysis*, a process involving both deductive and inductive analysis, wherein a researcher 'draw[s] on existing research and theory to inform the ... identification of emergent categories' (Renninger & Bachrach, 2015, p. 63). That is, after transcribing each recording, the researcher engaged in close reading of the transcripts, which were first coded

according to topics derived from a review of previous literature (*deductive analysis*), including:

- (a) The types of informal practices that learners participated in
- (b) How they felt and what they thought about when engaging in these activities
- (c) Their reasons for engaging in them
- (d) Their attitudes towards these practices in general and as tools for language learning
- (e) Their general motivation for learning English

Subsequent analysis focused on identifying new emergent (sub)themes, as well as instances which further illustrated the nature of each of these categories (*inductive analysis*). For example, the data were used to explore how the different engagement types were represented in the students' informal L2 practices, and their relationship to the learners' motivation for language learning and perceptions of ISLL.

To demonstrate the consistency of the qualitative analysis, the final coding framework established by the researcher was used by a second coder (an English-German early-career SLA researcher) to code one of the interviews. Inter-rater reliability was assessed using the procedure set out by McAlister et al. (2017). The two coders were considered in agreement when the same code was applied to approximately the same excerpt of text. The rate of agreement was calculated by dividing the number of instances of agreement by the total number of codes. The initial rate of agreement was only 75%, mainly due to varying interpretations of some of the codes relating to the various dimensions of engagement. Through discussing the discrepancies, the two coders clarified the definitions of some codes and

established a final coding framework (Appendix C.1). In coding of one more interview using this adapted framework, the agreement between the coders reached 87% (more than 80% agreement is generally considered to be indicative of good reliability).

5. Findings Phase I: Qualitative Findings

This chapter presents the main qualitative findings from the semi-structured focus group interviews. It is divided into three sections: The first focuses on engagement, mapping the informal second language practices of the sampled students, their intentions in participating in these activities, and reflections on their concurrent thoughts and feelings. The second section contains a summary of the findings regarding the participants' general motivation for learning English as a second language. The last section discusses the students' attitudes towards informal and formal learning, primarily focusing on their perceptions of the effectiveness of these modes of learning and the types of knowledge gained from each.

5.1 Engagement: How and Why do These Students Engage in Informal Second Language Practices?

This section gives an overview of the different types of English-language activities that were most popular among the interviewed students, as well as their reasons for engaging — or not engaging — in them. It must be noted that the relative popularity of these activities was not strictly quantitatively determined, for example via questionnaire data, but instead inferred by the researcher from the interview transcripts. Obtaining counts of exactly how many students engaged in a certain activity proved difficult because, in addition to explicit mentions of an activity, many of the other students in the same focus groups simply agreed with accounts of their fellow participant (e.g. through muttering 'hmm' or 'yeah'). For this reason, simple quantificational phrases ('some', 'a few', 'most', etc.) are used throughout this chapter instead of precise counts, except in rare cases where there was no ambiguity in the data.

By far the most common informal activity among the pupils in this study was listening to music. Almost all participants reported listening to songs for several hours a day, overwhelmingly in English. Interestingly, their awareness of this as an English-language activity was limited: When asked where they encountered English outside of school, most participants first mentioned social media or online videos. When the researcher specifically enquired about music, many reacted with surprise at the realisation that they were listening to English for several hours a day through songs. It seems that, quite often, the students just played music in the background while engaging in other activities such as commuting, doing homework or chores or chatting with friends. Even so, most participants also described periods in which they focused intently on the music, often also while looking up the lyrics online in order to better understand their meaning or to be able to sing along.

After music, watching online videos (mostly on YouTube) was the second most popular activity. The participants watched a wide variety of genres, including how-to tutorials, educational content, videos about beauty and fashion (only popular among girls), video game ‘Let’s Plays’⁸, video blogs, prank videos, and other YouTube-typical formats. Some participants subscribed to particular English-language channels, that is, they regularly watched content uploaded predominantly by British or American creators. These students reflected that the videos of English-speaking creators were of higher quality, more novel, and creative than those made by German speakers, because the use of the English language allowed creators (even though they might not

⁸ Let’s Plays are a popular style of online video showcasing an individual player’s experience of a video game through screen capture and voiceover commentary, sometimes also including video recordings of the gamer’s face to show their reactions.

be native speakers themselves) to reach broader audiences. The pupils also liked learning about cultural differences between their own daily lives and those of young people living in the UK or the US through watching these videos.

What interests you⁹ is just almost always in English. I mean, you are usually not interested in the German ones, because they are just so ordinary. (Malena)¹⁰

I don't know, sometimes their videos are just more interesting. Or because you see the German ones so often. It's just something different. And their lives are also just... different. Even just seeing their surroundings, that's really nice. [...] How they live in America and all that, when they blog about it. And they just show you everything, what they do and everything, and that's great. (Constanze)

Most pupils, however, did not purposefully seek out English-language content on YouTube in this way. Rather, they reported searching for videos on a particular subject (e.g. a video game, a celebrity, a fashion trend) and finding mostly videos in English. Some also described being recommended English-language content by the websites themselves (i.e. YouTube's 'recommended videos' side-bar or similar functions on video subscription platforms like Netflix or Amazon Prime). This illustrates the role that algorithms can play in generally increasing learners' contact with English online:

[A]nd then you happen upon an English video, and then on the side there are always the recommended videos and you think 'Oh, that also looks interesting' and you click on it. And then you just keep going and keep going and at some point, you look at the clock and think 'Oh no, it's gotten pretty late! (Diana)

⁹ 'You' appears in many of these quotes as a translation of the German indefinite pronoun '*man*'. It is used when making a general statement to signal that it applies not to a particular person, but to a whole group of people. Presumably, the interviewed students were signalling that they thought their personal experiences were representative of those of their peers.

¹⁰ Pseudonyms have been given to all participants.

Besides YouTube, other popular platforms for watching English-language videos were Instagram and Snapchat. Notably, a majority of the girls but none of the boys in this sample used Instagram, where they mainly followed celebrities they knew from other media (e.g. actors, musicians). Much like on YouTube, some girls also followed fashion and lifestyle bloggers on Instagram, whom they found more interesting than German creators because they talked about trends or products that were not available or well-known in Germany. The girls described checking the app regularly throughout the day, scrolling through photographs, reading captions, and watching brief videos. Although they would use it for only a few minutes at a time, several girls estimated that they spent up to one hour cumulatively per day on Instagram. Snapchat was used in much the same way but by many fewer participants: It was primarily used by girls to watch videos posted by popular celebrities.

Surprisingly, other generally popular social media platforms, including Twitter and Facebook, were used (at least in English) by only a few students in this sample. Only five participants mentioned occasionally using Facebook messenger to have short, synchronous written exchanges with friends or extended family members from abroad. A mere two participants talked about their use of Twitter: One female participant used it to read posts by celebrities she liked, and one male participant engaged in discussions about US politics and Brexit negotiations in the UK.

The fourth most popular activity among the interviewed pupils was watching TV series and shows. By the researcher's best estimation, just under half of the participants regularly did this in English. Given that almost all British and American programmes on German television are dubbed (with the exception of those broadcast on English-language channels, such as BBC World and Comedy Central, which are only

available through premium TV packages), the participants mostly streamed English-language shows through web services, such as Netflix and Amazon Prime. In fact, streaming TV series online, or watching content produced by these streaming services, has become so popular that the vast majority of the interviewed participants (and German language users in general) have started referring to them simply as *serials* (German: 'Serien') instead of *TV series* (German: 'Fernsehserien'). This terminology will therefore be adapted throughout the rest of this thesis.

Even on streaming websites, a majority of programmes are available with German voice tracks, so that watching them in English shows a certain degree of intentionality. However, only one student reported watching serials in English specifically to improve her language skills. The rest cited three main reasons: First, if they were following a series of which new episodes were still being released, they would watch in English because they did not want to wait months for the release of the German version (and risk seeing 'spoilers' online).

Second, many students voiced an intrinsic preference for listening to the English language, which they perceived as somehow more beautiful or expressive than German:

I don't know, I just like the language better. Somehow, uh, it's — I don't know how to describe it. I just like it better, how they write and speak. Just like, the style. You get— I think that you notice a difference. (Marie)

Also I just like it more. The language alone, you can see, I think, more of the thoughts that the writer really had. And not just the bad German translation. (Nadja)

This was closely linked with the third reason for choosing original English-language content: A dissatisfaction with the quality of the German translations. A major aspect of this was a strong dislike, shared by a large proportion of the participants, for

hearing dubbed voiceover tracks, especially when they did not match what was shown on screen:

And in German, the dubbed voices are just sometimes so annoying! I mean, the lips always lag behind what they are saying. And then it feels like five actors, from five serials all have the same voices. [...] Or those shrill voices, that's just really exhausting. I mean, the dubbed voices for American movies are just so over the top. And that's kind of really strange. (Insa)

In addition to disliking the dubbed voices themselves, many participants thought that the translations did not successfully convey linguistically or culturally specific humour, sometimes left out vital pieces of information, and generally just 'sounded bad'.

Just because many translations— I mean, many phrases or something and jokes, if you can call it that, are just in English. They don't make sense if you can't really translate them. And in German sometimes not at all. I mean, some rhyme or something is totally different in German, a whole other word. They're entirely different phrases. (Rosa)

Sometimes there are things in English that you don't even notice in German. I don't know, that they just leave it out or something? [...] That they just roughly translate it, but then you [...] don't get half of it at all. And then, if you watch it in English right away, then... It's often much easier to understand! (Anna)

Two participants reported similar experiences with playing German translations of video games, which were originally released in English:

I also play a lot of single player computer games in English because of the bad [German] synchronisation. [...] I mean the dubbing is not great and therefore, uh... Most of the word play takes place in English. You can't translate it. (Christian)

Yeah, sometimes if you have the German translation, it barely makes sense, like the sentences. Then you kind of have to play it in English or you won't know what to do. (Tim)

Video games must also be mentioned as an activity that was quite popular, particularly among the boys interviewed in this study: All ten of them reported playing games in English, sometimes for up to three or four hours per day. By contrast, only seven of the girls (about 20%) mentioned playing video games at all, and only one of them said that she played regularly for several hours. The boys were also much more likely to connect with other gamers when they were playing: Six of them, compared to only one of the girls, mentioned that they regularly used text or voice chat to talk about strategies when playing multiplayer games. They were communicating in this way with people they had met in-game, who came from all over the world. That is, their interlocutors were not necessarily English native speakers, but English was the default language of communication among them.

Beyond the context of video games, only eleven participants (c. 20%) reported sometimes (between once per month and once per week) using English to communicate with others, either in writing or through speaking. All of them communicated with people they knew 'in real life' (i.e. not friends they met online), including relatives or family acquaintances who lived in other countries or friends they made on holidays abroad. Those students who did not have such personal connections only used English to communicate with others (besides in the classroom) if they travelled abroad.

Even though the internet is often lauded for affording language learners with many opportunities for productive language use, none of the students in this sample engaged in online discourse (with the exception of one boy who, as described above, used Twitter to join in political discussions). In fact, there seemed to be a certain social

stigma associated with engaging in conversations online (other than in direct messages with personal friends and acquaintances). This was apparent in all of the focus groups through the embarrassed laughter and judgmental tones that laced discussions around active participation in social media and using the internet to make new connections. Some participants also commented explicitly on this stigma, as well as their concerns about online privacy, which were often related:

I don't write comments in general. [laughs] I find it really embarrassing and uncomfortable. (Marie)

I just always think, no, that's a little odd. Because... I don't know why, but it's really embarrassing and uncomfortable to write like that online! Maybe if I'm tagging a friend in something funny. That's still okay. But, [...] I wouldn't write to strangers [embarrassed laughter from the other participants]. That's odd. (Lara)

I don't do it because, I don't know... I just feel that, no, other people can see it, too. Especially if— If you are commenting on something and other people follow you, too, then it always shows them what you said. That's embarrassing and I don't want that. [agreement from the other participants] So I don't tend to do it. (Insa)

Like writing and speaking, reading — beyond short texts the pupils encountered on social media — was generally not a popular activity in this sample. Some participants did remark that they sometimes read texts online in English, for example if they were searching for recipes or news articles. This kind of almost incidental (or at least less purposeful) contact with English on the internet (see also the paragraphs on YouTube videos above) also happened, but much more rarely, in the pupils' physical environments. For example, one girl reported frequently reading magazines in English, which were available to her at home because her mother had taken out a subscription. Another pupil was inspired by her sister, who had been an exchange student in the US, to watch TV shows and films with her and to read books in English.

In addition to this, only two more participants, who cited reading as one of their main hobbies, reported regularly reading books in English. Their reasons were much the same as for those participants who watched serials in English: They were unhappy with the quality of the German translations and/or they had an intrinsic preference for the English language itself:

I don't know, I just like the language better somehow. [...] I mean, I just like the way they write more. And their writing style and such... I think you can really see the difference. (Marie)

For the remaining participants, reading books in English was an activity they engaged in rarely, if ever. Even so, they seemed to think that they *should* be reading in English, or at least that this was the type of activity the researcher wanted to hear about (demonstrating positive response biases): For example, some participants made remarks such as 'once, I read a book in English' or 'I'm reading one at the moment'. Further probing, however, revealed that these were exceptions, most often because the student had been given an English book by someone else (e.g. parents, other relatives).

In part, this unpopularity of reading was because a majority of these students didn't read many books in any language. However, for those who did like reading, it seemed that it was often a bedtime activity and that they did not feel focused enough to read in English at that time (more below, Section 5.1.2). This also came up in the context of serials and films, as a reason for why some students preferred watching them in German. Finally, three students commented that, even though they understood most of the words in English books, they did not like to read them because they did not feel as immersed in the story as when they were reading in German:

And if you are reading a book in German, then you somehow picture it in your head like a film. Sometimes when I read books [...] you can't imagine it. I mean, you can't picture what's going on. That's not so good. That makes books a bit more difficult. (Christian)

Lastly, there were some other English-language activities in which only one or two of the participants engaged. These included listening to podcasts; reading newspapers, recipe books, or comic strips; programming; and acting in an English-language play.

As discussed above, there were some English-language activities (listening to music, watching online videos, serials, or films, browsing social media platforms, and playing video games) which were popular with large proportions of the sampled students and others in which only a few of them engaged (e.g. speaking or writing in English, reading books). Nevertheless, there was great interpersonal variation in the types of activities that each participant preferred, as well as in how much time they spent on them, by their own estimation. It should be noted, however, that the vast majority of the participants found it very difficult to estimate how much time they spent on informal English practices, per day or per week. On the one hand, they had trouble estimating how much time went into activities in which they engaged often but very briefly, such as checking social media or engaging in short written exchanges with friends. On the other hand, they also found it difficult to assess how much time they spent in prolonged informal contact with English when it was not their main focus, specifically when they played music or put on a YouTube video while also doing something else.

Without accounting for music, which most participants reported listening to for several hours per day, average estimates for the cumulative duration of informal practices ranged between 3 and 7 hours per week (30 minutes to 1 hour per day). For

most of these participants, engaging in the aforementioned activities — including watching serials and films, playing video games, connecting with others through social media, and reading — took up a much larger part of their free time, but they sometimes preferred to do so in German. This was especially true for participants who felt that they were not proficient enough in English to understand what they were reading or listening to, or had to concentrate so much to in order to reach adequate levels of comprehension that it detracted too much from their enjoyment (see Section 5.1.2).

The few pupils who spent significantly more time (up to 28 hours per week/ 4 hours per day) engaging in English-language informal practices than their peers were mostly boys who frequently played video games for several consecutive hours or girls who followed many English-language serials. At the lower end of the range, there were some students who estimated engaging with English for no longer than 2 hours per week (about 15 minutes per day) in their free time. They spent most of their free time on hobbies that did not involve media consumption (e.g. playing sports or music) and their only contact with English, besides through listening to music, was through online texts and social media.

It is important to realise that the amount of engagement in informal activities also varied within participants, not just from day-to-day, but across the year as well. According to the interviewees, fluctuations could depend, for example, on the kind of activities that they were most interested in at the time, how much free time they had available (e.g. exam periods versus weekends and holidays), but also on factors like the weather, since they were more likely to spend time outside with friends — away from serials, games, and social media — if it was sunny.

All in all, it can be concluded that there are vast inter- and intra-individual differences in informal second language practices, regarding the types of activities and time spent on them (both aspects of behavioural engagement). In the following sections, the question of *how* these students engage in ISLL will be explored further by considering the three other aspects of engagement that are part of this study's theoretical framework: affective, cognitive, and linguistic engagement.

5.1.1 Affective Engagement

Affective engagement was defined in this study as the learners' emotional states while engaging in ISL practices, including their levels of interest and enjoyment and how much they liked the activities. By and large, the interviews revealed that affective engagement was the main driving force behind ISLL: Over and over, the pupils stressed that they engaged in informal practices mainly for entertainment purposes, so they chose activities which they knew they would enjoy and which aligned with their personal interests.

I think it just depends on your interests. Because some people prefer reading books to watching serials, or something like that. But most of the time, if those are your interests, then you do it of your own accord. (Eileen)

I mean, the content is interesting to you and also the people. If you watch a video, then you like the person or what they do. But that it's in English... It's not like you want to learn English and then you watch it, or something like that. (Jennifer)

Interest and enjoyment were not just important motivators for the participants to engage in informal activities in the first place, but they also seemed to be the main determinants of participants' willingness to invest mental effort in an activity and to persist in it even if it required closer attention (cognitive engagement, Section 5.1.2):

I once tried watching a series that— I already watched it all the way through in German and I tried to start it in English. Because I thought maybe you can learn a bit more and you already understand what happens. But I didn't continue it, because it wasn't as much fun. (Karin)

For example, on Snapchat there are sometimes those articles. I tried reading them sometimes [...] But then I didn't understand what they were writing there at all. And then I thought 'Well, somehow it doesn't really interest me anymore'. And then I just left it. (Johanna)

Similarly, interest also seemed to be essential in determining the extent to which the pupils were willing to employ language-related strategies to increase their comprehension in informal activities (linguistic engagement, Section 5.1.3):

It generally depends on [...] whether the video really interests me or whether I just clicked on it for fun and don't really need to see it. Otherwise I would just make it go away, I'd say. But if it really interests me, I would maybe rewind and watch it again, or look something up if I didn't understand it. (Lisa)

It really depends on the video. When it's one that I'm really interested in, in the topic, then I would maybe watch it again to understand it better. But otherwise, if it's not that important, I wouldn't bother, I think. (Finja)

Sometimes, a more conscious linguistic focus, borne from an intrinsic wish to engage with the language, could also be the very thing that piqued or sustained a participant's interest and enjoyment. In these (comparably rare) cases, the relationship between linguistic and affective engagement seemed to be mutually reinforcing:

This show, Thirteen Reasons Why. I first watched it in German and then in English. Because I liked the show, and it would have been boring if I watched it again in German, I decided to watch it in English. [As] a challenge! (Johanna)

All in all, these results seemed to indicate that there might be some sort of threshold of interest and potential for enjoyment which an informal English-language activity

has to cross in the participants' mind to be worth doing (in addition to being fairly easy to understand, see Section 5.1.2). Furthermore, while already engaging in an activity, higher levels of affective engagement seem to be associated with an increased willingness to invest mental effort in the activity at hand (cognitive engagement) and focus on the features of the language contained within (linguistic engagement).

5.1.2 Cognitive Engagement

This dimension of engagement describes the learner's cognitive state during ISLPs, including the extent to which they were focusing on the activity itself (vs. extraneous stimuli) and on comprehending of the *meaning* of what they were processing. This factor was found, in addition to affective engagement, to play an essential role in determining whether or not individuals chose to participate in certain informal activities. That is, the participants who *did* very frequently engage in ISLPs often commented on the fact that it took little conscious effort for them to understand the meaning in these activities. Some even claimed to perceive no difference between watching a video, reading a text, etc. in English and German.

I mean it's just really practical and really pleasant to watch a series in English, also because you perceive it just like it is in German. It just feels normal. (Rosa)

For me it really doesn't matter if it's in English or German. (Robert)

I would say you don't really think about [the fact] that it's in English. It's just... It's just really normal. It's normal. You just do it. (Insa)

By contrast, some participants who did *not* engage in ISLPs as often explained that this was because they had to concentrate too much to understand what was going for it to be an enjoyable experience:

But when I just want to do something to relax, I wouldn't want to suffer through the English to play a video game. I would just do it in German, so that I can just do it without having to try so hard — I mean, to understand what it's saying and what I have to do. (Jill)

It's sometimes less fun because you have to concentrate more to understand it in English. Otherwise it's always like... You just look at it and understand everything. (Karin)

In the previous section, it was observed that there seemed to be a positive association between affective and cognitive engagement, with participants being willing to put more mental effort into an activity in which they were interested and which they enjoyed. However, the quotes above show that there can also be a negative relationship in some cases, where the higher required levels of cognitive effort take away from enjoying the ISL activities.

How much focus the participants felt was required did not just vary between different students (it seems likely that this would be related to variations in proficiency), but also according to some features of the activities themselves. For example, many participants reported struggling with fast rates of speech (in serials and YouTube videos) and with reading books or watching serials in English which were narratively more complex:

For example, in serials which are really difficult... Like I said, I really like Sherlock. But even in German that's sometimes really hard to follow. And then when I even just watch the trailer in English, it's really hard for me to understand what's going on. And then he also talks so quickly—I think that's stupid. I want to be able to understand for myself how he got to his conclusions. (Lisanne)

Generally, the students agreed that some types of activities inherently required more attention than others. In particular, a large number of participants independently

compared watching serials to reading books, but there was disagreement over which was easier:

I think with serials it's easier in English, because with books you have to somehow read it. And with videos it just... It all just comes at you. And I think that's easier. When I read books, I sometimes skip some sentences. I just go over the pages and then I forget what is going on. (Eileen)

I think when you're reading you understand more than when you're watching it. Because you can see it written down. And you don't have the pressure, that you have to understand it right away. Because with a series, or a YouTube video, it just keeps on going. (Monika)

All things considered it seemed that, similar to affective engagement, the level of required cognitive engagement played a role both in determining whether or not a pupil would (continue to) participate in an informal activity. How much focus was required varied greatly between different students and might be related to language proficiency. The data also showed that, if the required level of cognitive engagement was too high, this could negatively affect students' enjoyment of informal activities (affective engagement) and discourage them from seeking out further informal L2 contact. Nevertheless (as discussed in the previous section) there also seemed to be a positive relationship between affective and cognitive engagement, in that students were more willing to invest effort into understanding an activity which they found highly enjoyable and interesting.

5.1.3 Linguistic Engagement

The final aspect of engagement considered in this study was linguistic engagement, that is, the extent to which the participants focused on processing linguistic features and improving their language skills while engaging in ISLPs. In general, the

interviewed pupils did not focus much on the language in informal activities, as illustrated by statements such as this:

To be honest, I don't really pay attention to the language when I'm listening to music or watching a serial or something. Because, for example, I'm more interested in the music than what they are saying exactly. (Eileen)

However, a substantial number of pupils also described how sometimes their attention might be caught, in the moment, by a particular expression or linguistic feature, which they would then attend to. Most often, this was a word or phrase that they felt was essential to understanding the overall meaning, or one which appeared repeatedly (e.g. in the chorus of a song or throughout important scenes of a TV show).

Sometimes it's unimportant information and sometimes there are key terms in there, that you have to understand, that come again and again. And if you don't understand those [...] then maybe you look them up or something. (Lisa)

If the participants decided to employ language-related strategies, this was almost always to increase their understanding of the overall meaning, instead of with the intention of learning new language features:

Monika: But if you hear something over and over again that you didn't understand, and then you look at it again, then I do think that you hear the language more. Because you would like to understand it. [...] I think then you pay more attention.

Interviewer: But then is that about understanding what's going on, or about learning words or phrases, or some grammatical structure?

Monika: No! [laughs] It's definitely just about what's going on, generally.

When describing how they dealt with comprehension issues, without exception, the participants said that they would first try to derive the meaning of what they did not understand from context, either the surrounding text or any available visual cues (for example in videos). If this failed, they described using other strategies, including rewinding a video or re-reading earlier parts of a text, using online dictionaries or translation apps, turning on captions (for films and serials), looking up lyrics to songs, and finding a German translation (e.g. of a song, text, or film scene) and comparing it to the English version.

Only three pupils described employing strategies expressly to *learn* new language features they encountered: Two participants told the researcher that they tried to mentally review new words or phrases they discovered (e.g. at the end of a TV show episode) and to think about other situations in which they could use them in order to increase their chance of recalling them in future. Moreover, one participant said that she kept a list of words and expressions she had learned in her free time on her phone, which she reviewed before exams in the hope of increasing her vocabulary.

Speaking to the relationship between linguistic engagement and the other engagement dimensions, the participants generally agreed that thinking about the language could take away from their enjoyment and even their overall comprehension, especially during very meaning-focused activities such as watching serials.

Yes, but then the action just continues and then you miss something again and... I think you won't really understand what's going on and enjoy it, if you think about it so much. (Diana)

In English, if you have to think about the language, then you can't really understand the plot, because you have to think so much about what they even said. (Sarah)

This implies a generally negative effect of linguistic engagement on both cognitive and affective engagement. Nevertheless, as said before (Section 5.1.1), there was also a positive association between these dimensions to some extent, with participants being willing to pay more conscious attention to the language when they were focused on and enjoying an activity.

All in all, the focus group findings shed some light on the roles that the different aspects of engagement play in informal language learning. It seems clear that the different dimensions share close positive as well as negative relationships and that they can vary depending on a host of factors, including characteristics of the individual learners (such as their personal interests and their proficiency) as well as the activities themselves (including narrative complexity of the text or video and the rate of speech).

The next section will summarise what could be gleaned from the focus group interviews about the participants' motivation for language learning.

5.2 Motivation for Language Learning

Over the course of the interviews, many of the pupils revealed information about their general motivation for language learning by talking about why they thought knowing English was (or was not) useful to them currently and the ways in which it might benefit them in future. Remarkably, all of their reasons could be mapped onto the six

language learning orientations commonly cited in the SLA literature (Section 2.3), which will be discussed below in order of the frequency with which they were mentioned.

1. Travel orientation and 2. International orientation

For a vast majority of the participants, the language's communicative potential was the main reason why they wanted to be proficient in English. For them, the most important benefit of knowing English was its usefulness when travelling to other countries, or potentially living abroad in future.

People speak English so that they can be understood. I mean, that's the same in many countries, not just where they speak English. For example, we are also learning French. But if we fly to Spain or something, then I think we would also speak English. Because almost everybody will understand! (Lara)

Quotes like this illustrate that English has a firm place as a *lingua franca* in the mind of these students. Moreover, they did not recognise the same communicative potential in other languages, despite the fact that all of them were also learning at least one other language (usually French or Spanish) at school.

I would say, nobody needs French! [laughter from the other participants] But that's how it is! I mean, you need English! [...] English is the world language. Everybody has to learn English and that's great! (Sina)

The ways in which the pupils described being able to use English to communicate with others more closely mapped onto what the literature calls *travel orientation* (wanting to know English in order to travel or live abroad) rather than *international orientation* (wanting to know English to be able to interact with people from other cultural and

language backgrounds). That is, the students mostly mentioned that the language was relevant to them when they were abroad on holiday. Most of them did not find much use for it at home, in their daily lives, to connect with people from international backgrounds, unless they had personal acquaintances with whom they communicated in English. A few participants did note, however, that they imagined this might be different if they lived in a more urban area:

I also think here, it's not as important to know English well. But I think if you were living in a bigger city, like Berlin, then I think it would be pretty important to know English, because that's very international. (Philip)

In bigger cities, perhaps, because it's the language that everyone speaks — or should speak [laughs]. And then you can speak with others, no matter which country they come from. I mean, you can just make social connections beyond your own social group as well. (Tanja)

Altogether, it seems that the distinction between these two aspects of language learning motivation, travel and international orientation, is not particularly clear in this group of students. Because they came from a comparably rural town with a mostly heterogeneous German population, most of them were only able to make use of the English language's communicative potential — which they did recognise in theory and saw as a potential future benefit — when they were traveling abroad.

3. Instrumental benefits

The second most frequently mentioned motivation for learning English were *instrumental benefits*, that is, the advantages that knowing this language could offer in terms of future career and education:

I think that English is more and more a basic requirement. [...] Today it's everywhere. People our age have to learn it in school and if you want to study something later, you definitely need it. (Monika)

Of course, it depends on what kind of a career you choose. But it's very possible that you have to use English in your later life. (Janine)

Although most of the interviewed students did not yet have a clear idea of the kind of work or study they wanted to pursue after finishing secondary school, most of them seemed to be under the impression that English would play an important role in their future career. To contextualise these perceptions, it bears remembering that the participants in this study were all students at a *Gymnasium* (selective-entry school). There is a high expectation for graduates from this, the most advanced type of secondary school in Germany, to attend university and to later have a successful career. Instrumental benefits of English may not be such a strong motivator in other populations of German pupils.

4. Knowledge orientation

A number of quotes by different participants aligned closely with the so-called *knowledge orientation*: wanting to know English to pursue personal interests and to help in acquiring new knowledge. Statements like those below illustrate that these pupils recognised the value of the English language not just as a tool that would allow them to communicate with other English speakers, but also as a way to pursue their personal interests, access information, and develop new ideas.

I think, to pursue your personal interests... For example, the majority of pupils are probably not— [uncomfortable laughter] Not interested in the Roman and Greek gods, like me. So, in class you probably want to target the more general interests. But I can just go and read about it, and the more I read the more I can understand and learn. (Jennifer)

English just gives you a lot of possibilities. Not just for later in life, but also now. So that I can, for example, watch videos or go through English websites to learn about what all is out there. (Robert)

It must be acknowledged, of course, that there is a large theoretical overlap between the potential uses of English encapsulated by the knowledge orientation and the informal language practices that were the main subject of these interviews. That is, there is a possibility that these participants were biased, through discussing in the focus groups the many ways in which they used English in their daily lives, to perceive the language as more relevant to their personal interests and pursuits of knowledge than they may have under normal circumstances. Nevertheless, it is interesting that the participants often talked about these benefits without being prompted (i.e. before being asked explicitly about their language learning motivation), and that the language they used so closely resembled that which occurs in questionnaires used to capture this language learning orientation (cf. Section 6.1.2).

5. Intrinsic motivation

As mentioned previously, some students reported engaging in informal practices in English due to an intrinsic preference for the language, which they perceived as ‘more expressive’ or ‘more beautiful’ than German. Similarly, some participants also seemed to demonstrate a high intrinsic motivation for learning English, stressing that they wanted to learn the language because they liked it and because they enjoyed the process of learning languages more generally:

It's a bit of a challenge sometimes but I just think it's really fun! I like English! (Sina)

Just in general, I find it really interesting to be able to speak other languages. [...] It might be a bit annoying when you always have to learn vocabulary and stuff, but it's also interesting and like, good, if you can speak more than one language. (Helen)

6. Instrumental obligation

Finally, only two of the interviewed students displayed a clear *instrumental obligation* towards learning English, that is, they declared that they did not think that they particularly benefitted from learning English and were learning it only because they had to pass their English classes:

I mean, we have English classes which force you to study English. I would say without those, I probably wouldn't.... I don't think I really need English. (Jill)

It is interesting to note that these same pupils also reported that they did not like engaging with English informally because of comprehension difficulties but would sometimes try to do so specifically with an intention to improve their language skills.

5.3 Attitudes Towards Language Learning

One of the goals of the focus group interviews was to explore the participants' attitudes towards informal and formal language learning. In the following sections, the students' views pertaining to three themes will be discussed: the general effectiveness of informal and formal learning, the types of knowledge that the pupils felt could be gained from each, and differences in the affective states that the students associated with each context.

5.3.1 Perceived Learning Gains

The vast majority of the pupils agreed that most of their current English proficiency had come from school, with estimates ranging between 60 and 90%. Conversely, this does mean that all students also thought that at least a small part of their current knowledge had come from ISLL. The participants mostly credited their school classes with helping them reach a ‘basic level’ of English, which then allowed them to go out and use English in their free time, if they wished to do so.

I think 75% [comes] from school and 25% from what you do yourself. You need the basics first, to understand it. (Lisa)

Grammar and syntax mostly from school, yes. But once you have that... I think vocabulary alone.... 40% from class and 60% from serials, films, and some conversations with some people, everything else that's going on. (Rosa)

As the second quote illustrates, many participants attributed most of their knowledge of English grammar and syntax to formal learning, whereas they felt that they had acquired much of their vocabulary, beyond the foundations, in their free time (see Section 5.3.2). This was also why a number of participants stressed that, while they had gained most of their *total* English proficiency in school, informal learning played a much bigger role for them more recently, since they had become more advanced learners:

It's not a huge part overall. I think we did learn the majority in school. Because we already started in primary school. But then you don't use it outside [of class]. And... Now, there is a big part that's come from that, I would say, from serials or Instagram, that we learned on top of that. (Janine)

I think you learn more in school when you are still learning grammar. But after that... It's just about learning new words and broadening your vocabulary. I often feel that I just, when I'm at home— I feel that I learn twice as much as in my English lessons. (Marlon)

Inversely, these pupils also perceived their recent classroom activities as less valuable: They felt that a lot of class time was spent on introducing new ‘themes’, and with them, vocabulary that felt ‘far-fetched’ and ‘not relevant’ to their current interests or daily needs.

Because I think that recently — in the last two years or so — we didn’t really learn important vocabulary anymore, or anything new really. We just had new themes, that we just carried on with, which weren’t really relevant most of the time. But we didn’t really learn all that much new stuff. So, then I would say that I learned more in my free time. And not so much in class. (Lara)

In general, the degree to which the participants felt informal learning was effective also greatly depended on the language used to talk about it during the interviews: If asked whether they thought that they ‘learned English’ or ‘learned something new’ (in German: ‘lernen’) through these activities, most participants would answer negatively. This may be because the verb ‘lernen’ has a dual meaning in German; it is equivalent to both the English words ‘to learn’ and ‘to study’. The interviewed participants seemed to mainly interpret it in the latter sense. However, the vast majority did agree that such activities could help them, for example, to ‘improve’ (‘verbessern’) or ‘practice’ (‘üben’) their English:

It’s more of a matter of practice, rather than learning something new. (Eileen)

You’re automatizing it. Practice makes perfect! (Robert)

I would rather say that it helps to strengthen it, I mean, that you kind of revise it in a way, and then you can do it a little better. But not necessarily that you learn a lot of new things. (Lara)

But it’s not like you learn it from scratch. Actually, you already know, with songs for example, most of what you’re hearing. But I think it rather helps to reinforce it, when you hear it again in your free time. (Jennifer)

I do think it can be useful, if you learn the language through more than just engaging with it in class. In daily life, it becomes more and more ingrained. (Aaron)

In the literature review (Section 2.1), it was mentioned that ISLL is often compared to incidental language learning. Some participants, quite aptly, described learning through informal practices in the same way:

I mean, I'm not really learning directly in my free time. I'm more learning subconsciously. (Christian)

But in your free time, you just pick it up subconsciously and just, take it with you along the way. (Janine)

Yes, it's just sort of.... A side effect, learning something. On the one hand you understand the plot, but on the other hand you maybe get some other words or something, that you can then introduce in English class, too. (Jennifer)

Some participants believed very strongly that informal practices helped them to improve their English, but the vast majority of them could not name examples of specific words, phrases or constructions that they had learned in their free time. However, not all students believed so thoroughly in the effectiveness of ISLL, and even among those who did, it was highly dependent on the type of activity being reflected upon:

I would say, especially if you are reading English books, I think then you learn from it. But if you only hear it, I don't think you learn that much. (Diana)

I would say that you learn when you are watching TV, but I would also say that you probably learn more when you are reading. (Johanna)

Most participants thought that speaking more English in their free time would, in theory, be most beneficial for them; yet, this was among the least popular informal activities overall. Generally, the students also seemed to agree that reading more

books in English may help them to improve their language skills, but most of them also did not do this. When it came to the effectiveness of the activities with which many of them *did* actually engage regularly, opinions were more split: Most pupils did not think that watching serials and YouTube videos provided many benefits, although some thought it may help them to improve their listening comprehension. Even more participants expressed doubts over the effectiveness of Instagram and other social media as language learning tools, because they thought the language they encountered there was ‘too simple’, that there ‘was not enough language involved’, or that it was ‘not relevant’ to what they needed to know (for school).

The last point illustrates that there also were differences in how participants judged their language progress, depending on whether they perceived ‘English’ as referring to the language itself, or to the school subject: Those participants who stressed, elsewhere in the interviews, that English was important for them because they wanted to be able to communicate through it were usually more positive about the effectiveness of informal learning than those who judged their language proficiency in the way they were used to from school (e.g. exam grades, grammatical accuracy).

Well, I do think that sometimes it is the case that you think that it helps you learn. Especially — I don't mean in relation to school, but that some things support your English, to communicate and understand it better.
(Fabian)

Finally, a small number of participants mentioned that, while they did not necessarily think that they gained much new knowledge from informal second language practices, they valued them as opportunities to immerse themselves in the English language.

This, they felt, helped them to ‘get used to using [English]’, to get into an ‘English mindset’ and made them feel more confident and motivated to learn the language.

I always do it before we write a test or something. I might watch a film or something and I switch my phone to English. Just to get into it and so that I can read it a little. It helps me to get into the right frame of mind and makes me want to learn more. (Jill)

In most groups, being asked whether or not they thought informal language practices had contributed to their English proficiency also prompted the participants to consider the types of skills or knowledge that they thought were being advanced in school and in their free time. The outcomes of these discussions will be summarised in the next section.

5.3.2 Types of Knowledge Gained

Similar to the findings in previous research on ISLL (e.g. Kusyk & Sockett, 2012; Sockett, 2014b; Toffoli & Sockett, 2010), the pupils interviewed here thought that informal second language practices mainly promoted vocabulary learning. As noted in the previous section, many participants felt that, while they had learned most of their basic vocabulary (as well as all the necessary grammar and syntax) in school, what they had been taught in class more recently did feel relevant to their own and interests and needs.

I think the vocabulary that we learn in school, I already know that I will use it only once or twice, if at all, in my life. And if not, I could just describe it in a different way, I think. And when I'm watching a video or something, then it's something that really interests me, that I can use. (Rosa)

Similarly, many other participants also reported that informal practices helped them acquire a broader, yet more individualised vocabulary, since they could choose to engage with contents that was interesting and relevant to them. Furthermore, many also thought that an important benefit of ISLPs was that they exposed them to more ‘authentic’ English, as opposed to the ‘school English’ that they encountered in the classroom:

Because it's not just this school English, but like how they actually speak it. And it's like... youth language as well, in part. Or more commonplace language, that you wouldn't learn from your textbook. (Marie)

I have the feeling that you learn vocabulary in class that you can't really use in daily life. And in serials, [you encounter] more words that you can really use every day, and how they actually talk, and stuff like that. (Lara)

In this way, these pupils thought, informal practices helped them to acquire a more colloquial (less academic) vocabulary, as well as idioms and expressions which they viewed as an essential part of learning how the language is ‘actually spoken’.

I think it's like this, you learn all of the grammar in class. But how you get by in daily life, so to speak— You can't learn that in class. That's what I like, for example, when I use it in my free time. (Lisanne)

In addition to stressing that the types of vocabulary they were learning in their free time felt more relevant, many participants also explained that it was easier to learn new words informally because they were repeated often and presented in context, whereas in school they were only taught once, often in a decontextualized manner (e.g. from vocabulary lists).

The more you watch a series and the more an expression gets used, the more you understand it. Of course, it's the same in German: If you have never heard an expression, then you might react to it, but not understand it fully. And if you hear it more often, in context, then it starts to click. (Rosa)

And there's actions, realistic actions or at least situations how they could have been. [...] And then you immediately have an exact use, like an example that you can draw on if you're trying to somehow remember this word or use it. (Robert)

Some participants also mentioned that they felt the knowledge they gained informally was 'longer lasting': For school, they said they would often try to cram before exams but wouldn't remember much of it afterwards, because they never really got to apply the knowledge. By contrast, as mentioned in the previous section, the students felt that informal practices gave them a chance to apply and practise what they learned in school.

It's like... In the lessons you get the grammar and vocabulary and in your free time, when you use English, then you learn how to speak. [...] Otherwise you just have the pure grammar and all that and then, when you use it in your free time, you practise applying these things. (Nicole)

However, not all participants had quite such positive opinions about the outcomes of informal second language learning. For one, some students emphasised strongly that they thought learning the more colloquial, 'everyday language' — which their peers felt they could gain only from engaging with English informally — was completely unnecessary for them. In a way, these participants seemed to demonstrate a more 'external locus of control' (cf. Benson, 2011, Section 2.1) and strong positive beliefs in formal learning, trusting that the curriculum/teachers would provide them with all the necessary knowledge.

I think everything that you really need for later [...], you tend to learn in school. Because I think you wouldn't really need the more colloquial language, even if you are living there, I would say. And I think that in school, what you learn there, completely suffices. I don't think you need colloquial language. [...] I think it's enough to just be able to talk normally in English. [laughs] (Lisa)

Furthermore, the effectiveness of informal learning was not judged positively with regard to all types of learning outcomes. For example, a vast majority of the interviewed pupils did not believe that the types of informal practices in which they engaged frequently (which mostly involved receptive language processing) had a positive effect on their *productive* language skills (writing and speaking), to which they assigned the most value.

I just think, if I'm listening to it or looking at it, it doesn't stick. I mean, I understand it, when I hear it, but I wouldn't be able to speak it. (Jill)

The vast books or other longer texts; not on social media) could help them improve their reading comprehension, fluency, and speed. Similarly, listening to English (particularly by watching serials and online videos) was widely thought to have a positive effect on general listening comprehension, the ability to understand different accents, and even on their own pronunciation.

Altogether, the pupils interviewed in this study were most convinced of the positive effect of informal second language learning on their vocabulary knowledge. There was more disagreement about other specific outcomes of engaging in ISLPs, but most of the participants did believe that these activities in general could teach them to use and understand 'more authentic' English, improve their ability to understand the language, and boost their general confidence as L2 users.

5.3.3 (In)Formal Learning and Affective States

Most of the students' attitudes towards informal and formal learning related to their relative effectiveness for supporting various English language skills. However, a couple of participants also made interesting observations about the way that they *felt* differently in informal and formal settings, which will be described here.

Many of the participants who held positive views towards informal learning agreed that such activities could provide less stressful contexts in which they could practise their English, as compared to the classroom. For example, many students admitted that when they were speaking or writing English during their lessons, they were afraid of the judgment of their teachers and, to a lesser extent, their classmates. They reported that the teachers constantly paid close attention to their students' use of grammar and syntax and would correct them immediately whenever they made a mistake. This made many of them hesitant to speak up in class. However, those who did use English in their free time to communicate with other language users, online or in person, felt that they were able to communicate more freely in this context. This was because they felt able to focus solely on conveying the meaning they were trying to get across and to experiment with the language they knew without having to pay close attention to grammar and syntax.

Yes, that's exactly it, [nobody] corrects you. [other participants agree] You can just try and describe it. And you can understand each other with fewer words... And if the grammar doesn't fit quite right or something, then that's just the way it is. We don't care so much about that in German, either. [...] But mostly it's just that you are not under so much pressure in your free time. Or on social networking sites. If you are writing with friends, or something, you are not under as much pressure as in school. And then you just like thinking about it more. (Lara)

In other words, these students experienced less anxiety and more enjoyment related to language learning and use in the informal than the formal context. The most striking narrative of this kind was provided by Robert, who by his own estimation spent on average 3–4 hours per day playing video games and always concurrently communicated with other players via voice chat in English. In class, however, he said, he never spoke up, because he was too anxious and would start to stutter if he got nervous. The other students in the same focus group confirmed that Robert was among the most proficient English speakers in their class, but that he rarely participated in their lessons and seemed exceedingly nervous if he did.

Generally, it seems that it was not uncommon among the interviewed pupils to experience language-related anxiety during their English lessons and some felt more confident in their English use outside of school. Nevertheless, many students also stressed that they would not want to miss their English classes, mainly because they provided them with a certain incentive, or ‘pressure’, that they felt they needed to learn, particularly in order to acquire a basic proficiency in English grammar and syntax.

5.4 Key Findings

The most popular informal practices among the interviewed students were, in order: listening to music, mostly in the background while also doing other things, but sometimes as a focused activity; watching online videos on YouTube, Instagram or Snapchat; watching serials and (less frequently) films; and finally playing video games. The latter was particularly popular among the boys, who also often communicated with other players in speaking or writing while playing online multiplayer games. Using language productively, in speaking and writing, or reading

English-language texts (outside of social media) were among the least popular activities. Apart from holidays abroad, only students with personal connections to other (native or non-native) English-users would engage in speaking the language or using it for written communication.

By their own estimate, most of the participants spent between 30 minutes to 1 hour per day on informal second language practices. The full range of estimates went from as little as 15 min to as high as 4 hours per day. Altogether, there were vast inter- and intra-individual differences in informal second language practices, both regarding the types of activities and the time spent on them, that is, both aspects of behavioural engagement.

Regarding the other engagement dimensions, the interviews provided insight into the roles that these variables seem to play in informal language learning, that they share close positive as well as negative relationships, but also that they can vary depending on a host of factors, including characteristics of the individual students (such as their personal interests and their proficiency) as well as the activities themselves (including narrative complexity of the text or video and the rate of speech).

The participants' responses indicated that there might be some sort of threshold of interest and potential for enjoyment (affective engagement) that an informal English-language activity has to cross in the participants' mind to be worth doing, in addition to being fairly easily comprehensible (that is, not requiring too much cognitive engagement). How much focus the participants thought was required varied greatly according to the specific type of activity and between different students (most likely in relation to language proficiency). Furthermore, while already engaging in an activity, higher levels of affective engagement seemed to be associated with an

increased willingness to invest mental effort in the activity at hand (cognitive engagement) and to focus on the features of the language contained within (linguistic engagement). Conversely, the data also showed that, if the levels cognitive and linguistic engagement that are necessary for comprehension were too high, this could negatively affect pupils' enjoyment (affective engagement) and discourage them from seeking out further informal language use (behavioural engagement).

The interviewees were also questioned their general motivation for language learning, why they thought knowing English was (or was not) useful to them currently and the ways in which it might benefit them in future. The reasons that were most important to a majority of the students in this sample included being able to use English to communicate with speakers of other languages, especially when travelling abroad (international/travel orientation), providing them with advantages in their future career or studies (instrumental benefits), and allowing them to pursue personal interests, access information, and develop new ideas (knowledge orientation). By contrast, only a few students showed a clear intrinsic motivation (inherent interest in or enjoyment of the English language or language learning in general) or instrumental obligation (feeling pressured to learn English because it was a mandatory school subject).

Finally, the sampled students were asked about whether and how they thought informal practices (as opposed to formal classroom teaching and learning) could help them to improve their English proficiency. Most of them credited their school classes with being the source of nearly all of their grammatical and syntactic knowledge and helping them reach a 'basic level' of English, which then allowed them to go out and

use the language in their free time. By contrast, a great number of participants credited their informal practices with helping them to expand and individualise their vocabulary knowledge, improve their listening comprehension, and familiarise themselves with more authentic, colloquial registers than what they were taught in school. Using English informally also provided a less stressful context in which they could practice their English (listening and reading more so than speaking and writing), away from their teachers, who in their perception focused too much on grammatical accuracy, and the judgment of their classmates.

6. Methods Phase II: Instrument Development and Piloting

6.1 Initial Instrument Development

The sections below explain how the quantitative data collection instruments (C-tests, motivation and attitudes questionnaire, and informal language learning diary) were developed.

6.1.1 C-tests

C-tests (Raatz & Klein-Braley, 1981) are a form of gap-fill test (a variation of the cloze test [Oller & Conrad, 1971]) designed to serve as a simple, time-efficient, and easy-to-score measure of general language proficiency (Eckes & Grotjahn, 2006; Raatz & Klein-Braley, 2002). One test consists of several short texts in which the latter half of every second word has been deleted, starting with the second sentence. These words have to be reconstructed by the test-takers. The validity and reliability of C-tests has been thoroughly examined using a variety of techniques, including comparing test results across learner populations with different levels of proficiency (e.g. Coleman, 2002; Daller & Phelan, 2006; Reichert, Brunner, & Martin, 2014); criterion-referencing (e.g. Arras, Eckes, & Grotjahn, 2002; Babaii & Ansary, 2001; Hastings, 2002; Khodadady, 2014; Sarapuu & Alas, 2016); Exploratory and Confirmatory Factor Analyses (e.g. Hastings, 2002; Khodadady, 2014; Reichert et al., 2014; Sarapuu & Alas, 2016); and Rasch modelling (Eckes & Grotjahn, 2006; Lee-Ellis, 2009).

The bulk of extant studies seem to support a positive overall performance of C-tests as a tool for estimating the language proficiency of learners at various levels (Eckes & Grotjahn, 2006), although there is some disagreement about which specific *types* of language knowledge these tests predominantly measure. Specifically, it has been suggested that C-test scores are disproportionately influenced by learners'

lexical and morphological knowledge, while neglecting broader, communicative abilities (Alderson, 2002; Sigott, 2004). Yet, studies which have compared C-tests against other proficiency tests have generally found strong correlations both with measures of lexical and grammatical knowledge and broader tests of reading comprehension and written language ability, as well as weaker, but still significant links with speaking and listening skills (see review in Eckes & Grotjahn, 2006).

C-tests have also been shown to be capable of measuring language development, even over short periods of time. For example, Daller & Phelan (2006) found that a C-test could capture increases in the proficiency of French undergraduate students, who took an 11-week intensive course in English, with the same reliability as the widely used and thoroughly validated Test of English for International Communication (TOEIC).

However, researchers must take into account that small score increases may also reflect practice effects. Bogards and Raatz (2002), for example, administered a C-test to two different groups of participants within one week. In the intermediate time, one group received a brief (45 minute) training on taking C-tests, and the other group received no treatment. Even though the training did not significantly affect the participants' C-test scores as compared to the control group, both groups of learners showed increases in the range of 5-6%, which the researchers suggest represent a better familiarity with the test format. Similarly, Daller and Phelan (2006) suggested that score increases between 0.2–0.3 times the size of the standard deviation should be attributed to training effects.

C-tests are widely used as research instruments and placement tests at universities, preparatory colleges, and language schools, particularly in Germany. For this reason,

a norm-referenced English online C-test (onSET, *online-Spracheinstufungstest*; www.onset.de) was developed by the German TestDAF-Institute, which is also available for use in research. The onSET is intended primarily for testing university entrants — that is, learners who are at least three years older than the pupils in the present study, and therefore presumably more experienced language learners. According to Grotjahn (2002 – one of the leading experts on C-testing, and co-creator of the onSET), C-tests need to be customised for the population of language learners whose proficiency they are meant to measure, as content and vocabulary knowledge should be taken into account in the materials selection. Thus, the researcher decided against using the onSET in the current study and instead developed a new C-test suitable for the target population of year 10 pupils at *Gymnasien*. It must be acknowledged that this decision inevitably involved a trade-off between the increased targetedness of the new test and the established quality of the existing and widely used onSET. However, the researcher felt that the suitability of the test should take precedence and took care to extensively pilot the new tests.

Following the guidelines set out by Grotjahn (2002), two C-test versions were developed, to be used in counter-balanced order at the two data collection points in Phase III (see Section 8.2). The length of the tests was limited by the time that would be available for the participants to complete them (c. 20 minutes out of a 45-minute lesson). Grotjahn suggests that participants should be given five minutes per text, so the ultimate aim was to include four texts in each test. Grotjahn also recommends that more texts should be included in the initial versions of newly developed tests, so that some of them can be excluded after piloting if they produce statistically unexpected

scores or fit less well with the other items. Therefore, seven texts were selected for each initial C-test version.

According to Grotjahn (2002), the texts should not contain specialist vocabulary, require no specialist knowledge, be as authentic as possible, and also adequate in difficulty for the target group. He also recommends selecting the texts in collaboration with language teachers, who are familiar with the current level and learning history of the target group. For this study, short texts (or sections of longer texts) were selected from the website www.onestopenglish.com according to their level of difficulty ('elementary' to 'upper intermediate') and the aforementioned other criteria. This was done in collaboration with two English teachers whose classes had participated in Phase I.

25 gaps were created in each text by deleting the second half of every second word, starting with the second sentence. The texts were edited minimally for coherence, e.g. where context was missing as a result of excluding larger sections, or where the gaps led to ambiguities. Since Khodadady (2014) demonstrated that participants who are informed of the structure of C-tests (specifically, that exactly half of the letters are deleted from incomplete words) are more likely to indiscriminately produce wrong forms that do not fit within the semantic context, the instructions in the current study indicated only that 'part of the letters' were missing from some words (see Appendix B.1).

Both versions of the C-test were pre-tested with 20 native speakers of English (10 per version): According to Grotjahn (2002), texts for which native speakers' average scores fall below 90% should be excluded. This was not the case for any of the items in this study.

6.1.2 Motivation and Attitudes Questionnaire

In this study, motivation was operationalised as *language learning orientations* (LLOs; reasons for wanting to be proficient in English), as well as current and future *second language motivational selves* (L2MSs). The participants' attitudes towards language learning were operationalised in terms of their beliefs about the authenticity and effectiveness of informal and formal learning. Appendix B.2.1 contains a complete list of the motivation and attitude constructs included in the questionnaire, their working definitions, and the items through which they were measured.

New instruments were developed for measuring these constructs in the current study, rather than adopting any pre-existing questionnaire one in full. There were two main reasons for this decision: First, many items from extant questionnaires made specific reference to features of formal learning contexts (*e.g.* 'I'm always looking forward to my English classes'; Ryan, 2009) and thus were not suitable for investigating motivation in informal, out-of-school learning. Second, many instruments also contained items which, by the definitions in this study seemed to tap into different constructs from the ones for which they were designed, or multiple constructs at once (see Section 2.3.1). For example, Dörnyei and Taguchi's (2009) questionnaire on L2MS includes the item: 'Whenever I imagine my future career, I imagine myself using English'. This could be interpreted both as measuring the participants' desire to become successful future users of English (future L2 self) or tapping into their instrumental orientation (i.e. wanting to learn English to gain an advantage in one's future career). Thus, in the current study, items were adopted from existing questionnaires which (a) were as unambiguous as possible in representing the associated construct; and (b) did not explicitly reference the formal learning context.

Most of the items measuring LLOs and the L2MSs were adapted from existing questionnaires (Clément & Kruidenier, 1983; Dörnyei & Taguchi, 2009; Green, 1999; Hessel, 2011, 2015; Kormos et al., 2011; S. Ryan, 2009; Sundqvist, 2009). In addition, the findings from the focus group interviews (Phase I) were used to aid in the selection and, wherever necessary, the re-wording of items. Some new items were also written based on these findings, to complement those adapted from extant instruments and thus provide a more complete picture of the constructs in question. For example, one item was added to the knowledge orientation construct (defined as wanting to know English to become better educated or better able to pursue personal interests), which was intended to reflect a motivation mentioned by several frequent informal second language users in Phase I: *'I want to know English to access information that is not available or difficult to access in German (e.g. news, wikis & other online texts)'*.

Some new items were also added to the L2MS scales, which tended to emphasise the ability to *speak* fluently as an indicator of being a successful language user over other language skills (also noted by Hessel, 2015). To counterbalance this bias towards oral language production, several items were added which addressed other language skills such as reading and writing, or more generally communicating in or using the L2 (e.g. *'It's not a problem for me to use English'*).

All items were designed to be answered on ten-point rating scales. This format is sometimes recommended (over, for example, five- or seven-point scales) for its high internal consistency, reliability, and intuitiveness to respondents (e.g. Preston & Colman, 2000). The items addressing LLOs used a semantic differential scale of agreement (*'Don't agree at all'* – *'Completely agree'*). The current and future L2 selves constructs were both operationalised by the same items, but distinguished through

the use of two different rating scales: *'I desire this very much' – 'I don't desire this at all'* for the future L2 self, and *'This is already completely true of me' – 'I am still very far away from this'* for the current L2 self. This two-scale approach was adopted from Hessel (2015), because it allowed for a direct calculation of the difference between the two self-concepts, which is central to the L2 selves theory of motivation (see Section 2.3.1).

While there were numerous questionnaires from which items to measure language learning motivation could be adapted, the researcher could not identify any existing instrument for measuring attitudes towards both formal and informal language learning. Even though there are a number of questionnaires about language learner beliefs (e.g. Horowitz, 1988; Levine, 2003; Lightbown & Spada, 2013), all of them have been concerned mainly with formal language learning contexts, and some have focused specifically on particular features of formal learning, such as explicit grammar teaching and error correction (Loewen et al., 2009). Therefore, an instrument for comparing beliefs about informal and formal learning was newly developed for this study. The items were mostly based on views expressed by the focus group participants in Phase I (Section 5.3), although some items addressing attitudes towards formal learning were also adapted from the *Beliefs About Language Learning Inventory* (Horowitz, 1988) and a *Views About Language Learning* self-reflection tool by Lightbown and Spada (Lightbown & Spada, 2013). All items were designed to be answered using the same ten-point scale of agreement as the items targeting LLOs (*'Don't agree at all' – 'Completely agree'*).

All parts of the motivation and attitudes questionnaire were designed in English and subsequently translated by the researcher into German. Several measures were taken to check and improve the translation before piloting, following best-practice guidelines suggested by Dörnyei (2009a). First, the translation was checked for general comprehensibility by two German native speakers without specialist knowledge in Applied Linguistics. Where these reviewers noted issues with the translation, it was usually to point out that items were ambiguous in some way or did not 'sound natural' in German. In these cases, the researcher worked with the reviewers to find a suitable translation. Subsequently, an independent volunteer translator (English/German bilingual) with some subject knowledge in Applied Linguistics translated the German questionnaire back into English. The initial English version and back-translation were compared by both the researcher and the translator. It was agreed that the two versions closely corresponded, which supports the accuracy of the translation and is 'an indication that both instruments are asking the same questions' (Dörnyei, 2009a, p. 51).

6.1.3 Informal Language Learning Diary

The participants' engagement in ISLPs was operationalised in this study as comprising four components: behavioural engagement (the types of practices and the time spent on them), cognitive and affective engagement (the participants' thoughts and feelings while engaging in these practices), and linguistic engagement (the extent to which they focused on linguistic features and on improving their language skills). While the linguistic dimension has not been a part of previous engagement frameworks (see Section 2.2), it was included in the current study in the hope of increasing the applicability of the engagement framework to the specific context of

language learning. However, it must be acknowledged that the questions through which linguistic engagement was operationalised (see below) show some conceptual overlap with cognitive engagement, in that they were designed to determine the nature of the participants' attention during ISLPs. For this reason, the validity of linguistic engagement as a separate dimension of engagement was a particular focus for exploration both in the analysis of the qualitative responses (Phase I, Sections 5.1.2 and 5.1.3) and quantitative data (through Exploratory and Confirmatory Factor Analysis, Phase II and III, Sections 7.2 and 9.1.3.5).

A number of previous studies ISLL studies have used self-report questionnaires to gather data about various aspects of learners' informal practices. Usually, the participants were asked to estimate how often or for how long they engaged in various informal L2 activities *in general* (aggregated) or over a certain time span (retrospective), usually the previous week (e.g. Kuppens, 2007, 2010; Lai et al., 2015; Toffoli & Sockett, 2010). There are two major potential problems with these approaches: Firstly, learners may not be able to accurately recall their experiences retrospectively. This was a concern in designing the present study since it was found in Phase I that students may have had informal L2 contact frequently throughout the day without a high level of awareness: Many of the focus group participants, once they were prompted to think about it, expressed surprise at realising how much time they spent on ISLPs. Thus, it seemed likely that future participants might also have difficulty accurately recalling all of their L2 contact over a longer period of time. A second issue, which particularly concerns aggregated self-reports, is that participants may recall more recent experiences more clearly, and thus give them more 'weight' in their overall estimations. This was also a concern in designing the current study,

because many participants in Phase I remarked that their ISLPs could vary significantly from day to day, depending on external factors such as homework load and extracurricular activities.

As an alternative to retrospective or aggregated self-reports, other ISLL studies have adopted a language diary design, wherein the participants were asked to keep a daily record of their informal L2 contact, usually over the course of one week (e.g. Olsson, 2011; Sundqvist, 2009, 2011; Sundqvist & Sylvén, 2014). Although direct comparison studies of the accuracy of questionnaire and diary data in studies of language acquisition are lacking, there is some evidence from the wider Social Sciences to suggest that diaries can indeed provide a more accurate record of participants' experiences — and especially cognitive and emotional states — than one-off retrospective self-reports (e.g. Bonke, 2003; Conrath, Higgins, & McClean, 1983; Shiffman et al., 1997; Zunke et al., 2015). Of course, diary methods are not without drawbacks themselves: For example, there is a risk of 'reactance', that is participants may behave differently, or their perception may change, as a result of being asked to keep a record (Bolger, Davis, & Rafaeli, 2003). However, empirical studies of diary methods have found this effect to be minimal and short-lived, in most cases disappearing two to three days after the start of data collection due to habituation (Gleason, Bolger & Shrout, 2001, a.c.i. Bolger et al., 2003; Litt, Cooney, & Morse, 1998). After weighing the limitations of both methods and taking into account the aforementioned findings from Phase I, the researcher decided that a diary approach was likely to yield more reliable data in the context of this study.

The term 'diary' often evokes associations of long, free form writing about personal thoughts or emotional states. However, in research, diaries are often less open-ended,

consisting of self-report questionnaires administered repeatedly to examine ongoing events or experiences and their variability over time (Bolger et al., 2003). There are two main forms of diary designs suitable to the purposes of this study (e.g. Bolger et al., 2003; Wheeler & Reis, 1997): (a) interval designs, where participants provide reports at specific time points, e.g. once per day; and (b) event-contingent approaches, where participants make reports following an agreed-upon event, e.g. every time they log into Facebook.

While all of the aforementioned ISLL studies featuring diary methods have employed interval designs (requesting that participants report their activities once per day), an event-contingent approach was chosen for the current study for two reasons: First, as mentioned above, it was suspected that participants would find it difficult to recall, even at the end of one day, all of the ISLPs in which they had engaged. Second, unlike in previous studies, in addition to information about the participants' informal L2 activities (type, duration), this study also required participants to report on their concurrent cognitive and emotional states. Since it cannot be assumed that learners always reflect on their thoughts and feelings while engaging in ISLPs, these may be even harder to recall at a later time than factual information. An event-contingent design, however, prompts participants to actively reflect upon their experiences and record them immediately following the informal activity.

In this study, the diary took the form of an online survey (administered through www.qualtrics.com), which the participants completed outside of school, on their smartphone, tablet, or computer. In the focus groups, most students had expressed an explicit preference for a digital questionnaire over a pen-and-paper diary, as they always carried their phones with them but thought they were likely to lose paper

questionnaires or forget to carry them throughout the day. A further advantage of administering the diary online was that survey logic allowed the research to program the questionnaire in such a way that participants would see only the items relevant to the type of activity they were reporting at the time (e.g. when they were recording that they listened to music in English they were not asked follow-up questions about whether or not they used subtitles, as they would be for watching a film). Finally, because the entries were automatically time-stamped, the online survey allowed the researcher to get an idea of the participants' compliance (whether they filled in the survey after each activity, that is at various times throughout the day, or submitted entries in bulk at the end of the day/week). This information might also be used to get an idea of the changes in engagement over the course of the week during which the data collection took place. However, this variability analysis was beyond the scope of the current study, which aggregated the scores from all entries per participant for use in the statistical analyses (Section 9.1.3.5).

The participants were instructed to record all of their informal second language practices every day for one week, immediately following the first classroom visit. It would have been desirable to collect diary data over a longer period of time, particularly given that some of the participants in Phase I mentioned that their informal practices could vary greatly, e.g. from week to week or between school weeks and holiday periods. However, the researcher decided against this option because (a) the main focus of this study was not the change in ISL engagement over time, but its relationship to other variables, and (b) because keeping a diary for longer would have substantially increased the already high burden on the participants, which would have been a significant hindrance in recruiting a sample large enough to conduct the planned SEM analyses in Phase III.

The informal language learning diary included questions about the type of practices with which the participants engaged and for how long (behavioural engagement), as well as items targeting their cognitive, affective, and linguistic engagement during these activities (Appendix B.3.1). Each of these was designed to be answered on a semantic differential scale (e.g. *'All in all, this activity was [boring – entertaining]'*). In the online survey, these were represented by a slider scale, which allowed participants to visually select values between 1 and 100 but did not display the chosen numerical marker. This design was employed here — as opposed to the ten-point 'check box' scale used in the motivation and attitudes questionnaires — as the researcher felt it would allow participants to more intuitively express their thoughts and feelings during the ISL activities.¹¹

The cognitive and affective engagement items were, in part, adapted from existing instruments (Appleton et al., 2006; Cox, 2014; Egbert, 2003; Flowerday & Schraw, 2003; Fredericks, Blumenfeld, Friedel, & Paris, 2005; H.-G. Lee, 2012; Skinner, Kindermann, & Furrer, 2009). However, as with the motivation and attitudes questionnaire (discussed in the previous section), no one instrument could be adopted in full, as none was suitable for the context of this study: As mentioned in the literature review (Section 2.2), very few previous engagement studies have focused on a task-level (as opposed to a school- or subject-level) engagement. Out of these, most refer explicitly to formal learning contexts, and even fewer have investigated engagement in language learning specifically. Therefore, suitable items had to be picked from a range of instruments and, wherever necessary, reworded to suit the informal second language learning context of this study.

¹¹ Following the same reasoning, a 1–100 scale could have also been used in the motivation and attitudes instruments, but it would have been too difficult to implement in paper questionnaires.

Finally, since linguistic engagement is a new dimension of engagement proposed by the researcher, the items representing this construct were newly written for this study. As with the questionnaire for measuring attitudes towards (in)formal learning, these new items were primarily based on the qualitative findings from Phase I. The informal language learning diary was designed in English and translated to German following the same procedure as with the motivation and attitudes questionnaire, described in the previous section.

6.2 Piloting: Procedure

To pilot the C-tests, questionnaires, and informal language learning diary, they were completed by 104 students (35 boys, 69 girls¹²; ages 15-16) belonging to four different year 10 classes at two schools. 51 students (19 boys, 32 girls) were from two classes at one newly recruited school, while the other 53 (16 boys, 37 girls) belonged to the same two classes that had already participated in Phase I. 14 participants (13.2%) indicated that they had grown up bilingually. Of these, three were excluded from the analysis of the C-test data because they identified themselves as native speakers of English. Altogether, the remaining pupils had, at the time of the study, been learning English in school for an average of seven or eight years ($M = 7.7, SD = 1.05$). As in Phase I, the teachers estimated that these students had B1-B2 level proficiency according to the CEFR (Council of Europe, 2001).

The lesson set aside for the study began with a brief oral and written explanation of the general purpose of this research and a brief introduction to the format of the C-test, which was new to all students: The pupils were told that within

¹² For a speculative explanation of this gender imbalance, see Section 4.1.

the test, the texts increased in difficulty; that they should carefully read the titles and the first and last sentences of each text to get an idea of the subject matter; and that the gaps in the texts indicated where ‘parts of words’ were missing (but not that it was exactly half of the letters, to prevent the students from focusing overly on counting, cf. Khodadady, 2014).

While taking the C-test, the participants were reminded every five minutes that they should move on to another text if they felt stuck, and that they should not try to copy answers from one another. After they turned in their test papers, the students were given the rest of the lesson to complete the questionnaire. They were encouraged to ask the researcher if anything was unclear to them, and jot down any comments they might have on the questionnaire or test papers. The participants completed the C-test and questionnaires on paper; the researcher manually marked the tests and entered the data into IBM SPSS Statistics for analysis.

Over the seven days immediately following the classroom visit, the participants were asked to document their informal English practices, using the informal language learning diary. They were asked to save the survey to their mobile phones and to complete it every time they came into contact with English outside of school, through activities such as using social media sites, watching a film or serials, listening to music, etc. (Appendix B.3.2).

6.2.1 A Note on Participant Numbers and Test Time During the C-tests

Since one of the schools involved in the pilot phase teaches all their classes in double periods, more time was available there to conduct this study (90 minutes instead of 45 minutes). Therefore, to increase the number of responses per test version, the students at this school ($n = 54$ in two classes) were asked to take both versions of the

C-test. The order was counter-balanced, that is, one class completed Version 1 (V1) before Version 2 (V2) and vice versa. Altogether, there were 77 responses for V1 and 79 for V2 after exclusions (see above).

The first group of participants ($n = 26$) were given 35 minutes to complete the test: five minutes per text, as per Grotjahn's (2002) recommendation. However, after scoring and analysing this part of the data, the researcher realised that the students' scores were unexpectedly high (six out of 26 students scored above 85%). This seemed to indicate an emerging ceiling effect in the data, which would reduce the accuracy of the C-test as a measure of language proficiency. Therefore, the amount of time given for the completion of the C-tests was gradually reduced across the subsequent rounds of testing: The next two groups ($n = 52$, of which 26 completed both V1 and V2) were given 25 minutes (c. 3.5 minutes per text) to complete the test. The last group was given 20 minutes (c. 3 minutes per text).

It is worth noting at this point that, within each group, half of the participants took test V1 and half took V2, so that the reduction in test time did not affect the results of the two tests unequally. The researcher immediately scored the tests and calculated the descriptive statistics after each reduction (Table 6.1).

Table 6.1: C-test scores during Phase II, descriptive statistics: percentile scores by test time

<i>Test time</i>	<i>Number of students</i>	<i>Number of responses</i>	<i>Min</i>	<i>Max</i>	<i>Mean</i>	<i>Standard Deviation</i>	<i>Proportion of scores $\geq 85\%$</i>
<i>35 min</i>	25	25	42.3	96.3	71.4	16.3	23%
<i>25 min</i>	51	77	28.0	96.0	60.9	17.0	13%
<i>20 min</i>	27	54	34.0	94.3	62.8	16.6	15%
<i>Total</i>	103	156	28	96.3	63.3	17.03	17%

Finally, a one-way analysis of variance (ANOVA) was conducted to test the effect that the test time had on the overall mean scores. This ANOVA revealed a significant overall effect of the test time on the total scores ($F[2,152] = 4.15, p < .05, d = .40$). Planned contrasts revealed that having 20 or 25 minutes to complete the test significantly decreased the total scores compared to 35 minutes ($t[153] = 2.67, p < .05$ [1-tailed], $d = .43$). However, there was no significant difference between the scores of the participants who were given 20 and 25 minutes ($t[153] = .89, p = .19$ [1-tailed]). In the end, the researcher decided that the participants in Phase III should be given 20 minutes to complete the C-test instead of 25, so as to leave more time during the lesson to complete the motivation and attitudes questionnaire.

6.2.2 C-Test Scoring

The participants' C-tests were marked on paper and the scores (proportion of correct items per text) entered into SPSS for analysis. One point was awarded for each partially deleted word that was successfully reconstructed (25 per text, for a total maximum score of 175). Besides the exact word forms that had been in the original texts, from which the C-tests were constructed, some alternative answers were also awarded full points, if they fitted grammatically and semantically into the sentence as well as the text at large. The researcher determined which answers would be accepted with the help of two native English speakers (one British English and one American English speaker). Table 6.2 contains two illustrative examples of alternative answers that were/were not accepted. Finally, to account for partial knowledge, $\frac{1}{2}$ point was awarded for items that were spelled incorrectly, but which were recognisable and whose grammatical function was not altered (e.g. $\frac{1}{2}$ point was awarded for 'tatto' instead of 'tattoo', but not for 'spendd' instead of 'spentt').

Table 6.2: Illustrative examples of scoring alternate answers on the C-tests

Item	Original Text	Alternative Answer	Explanation
Accepted alternate answer (= 1 point)			
Version 1 Text 6	'For example, in big cities, the Blue Tit adjusts its song so that it can be heard more easily over traffic noise.'	'For example, in big cities, the Blue Tit adjusts its songs so that it can be heard more easily over traffic noise.'	The plural of 'song' was awarded one full point. Whereas in the original, the pronoun referent 'it' could be read as referring to 'song', in the alternative answer 'it' could refer to the 'Blue Tit'. Thus, the text as a whole is still grammatically and semantically acceptable.
Not accepted alternate answer (= 0 points)			
Version 1 Text 2	'This means falling over and pretending they were pushed.'	'This means * fallen over and pretending they were pushed.'	The simple past tense of '(to) fall' was not awarded a point because it does not grammatically fit into the given context.

6.2.3 Data Analysis

The data collected during Phase II was used to explore the measurement quality of the newly developed instruments. Generally, there are two main criteria by which quantitative instruments are assessed: *Reliability* and *validity*. Reliability is concerned with whether scores can be interpreted consistently in different contexts, for example over time (*test-retest reliability*), across multiple items (*internal consistency*), or across different researchers (*inter-rater reliability*) (Price, Jhangiani, & Chiang, 2015). Since the questionnaires and C-tests did not rely on the judgment of different researchers and the data could not be compared across time points (because the scores would be expected to change), the commonly used Cronbach's alpha was considered the best option for assessing the reliability in this study. Cronbach's alpha is a measure of internal consistency, which shows how closely related a set of items are. Generally, a Cronbach's alpha score of $> .7$ is considered sufficient, though smaller scores have also been accepted for instruments measuring psychological constructs or abilities (P. Kline, 2000).

Validity assesses the extent to which instruments actually capture the constructs they were intended to measure (Price et al., 2015). During the initial instrument development (Section 6.1), the items were chosen/designed to closely match the conceptual definitions of the constructs of interest. Ensuring that these items appear ‘at face value’ to measure these constructs and that they reflect all of their aspects is to consider the *face* and *content validity* of the instruments (ibid.). As is often a problem in social scientific research, there were no objective criteria for measuring the constructs in question (motivation, engagement and attitudes), so that it was not possible to inspect the instruments’ *criterion validity*. Instead, it was decided to focus on the *construct* and *discriminant validity* — that is the extent to which the proposed multi-item scales reflected a single underlying factor and did *not* reflect conceptually distinct but related constructs (e.g. affective vs. cognitive and linguistic engagement).

A common method for inspecting these types of validity is factor analysis, a multivariate statistical method which identifies clusters of variables by their shared variance (Field, 2018). There are two types of factor analysis: *Exploratory Factor Analysis* (EFA) and *Confirmatory Factor Analysis* (CFA). In EFA, researchers do not make assumptions about the factor structure underlying their data, instead exploring the inter-relationships of a set of item scores to investigate how many latent factors emerge, and which items most strongly represent them. This ‘bottom-up’ method is therefore most often used during the earlier stages of instrument development. In EFA, any set of indicators with sufficiently strong interrelationships are taken to load onto the same latent factor, so that the extracted factor models can sometimes change drastically with small shifts in the data. EFA, therefore, can only provide preliminary evidence of the factor structures underlying a set of measured indicators. CFA,

meanwhile, serves to test specific hypotheses about the number of latent factors in a set of items, which have been constructed based on theoretical knowledge and prior empirical research. That is, CFA is used to validate instruments by statistically testing how well factor structures specified by the researcher are represented in the available data. This 'top-down' method is therefore often used to check the validity of existing instruments in new contexts, e.g. when sampling from a new population or conducting research in a previously unexplored setting.

In this exploratory phase of the present study, EFA was therefore used to uncover the underlying factor structures of the proposed instruments. In this way it could be investigated, for example, whether the items aimed at measuring the learner-internal aspects of engagement (affective, cognitive, and linguistic engagement) actually measured multiple distinct constructs. By exploring which items most strongly represented each factor in the current data set, EFA was also employed to reduce the number of items per construct. To strengthen the claim for the validity of the instruments, the emerging factor structures were also tested using CFA on the larger data set collected in Phase III (see Section 9.1.3.5).

7. Findings Phase II: Evaluating and Finalising the Instruments

This chapter presents the results of the Phase II (pilot study) data analyses, which included checks for normality and outliers, reliability analyses, and Exploratory Factor Analyses to investigate the validity of the newly developed quantitative instruments and aid in the selection of the items which were most suitable for measuring the target constructs. The results are grouped by construct, starting with the proficiency data obtained by the C-tests, followed by the questionnaire data on motivation and attitudes, and finally the engagement data from the informal language learning diary.

7.1 C-tests

7.1.1 Descriptive Statistics and Equality of Means

First, for each of the two C-test versions, the overall and per item (i.e. per text) mean scores were investigated (Appendix D.1). This revealed that the total scores for V1 were slightly lower than for V2. Since the data stemmed from a partially overlapping sample (i.e. some participants completed both V1 and V2), a mixed model analysis was used to conduct a test of equal means. This revealed that the mean scores for V2 were significantly higher than those for V1 ($F[1,58] = 32.5, p < .001$). This difference seems to indicate that, overall, the two versions were not equal in difficulty. Less likely, but still plausible, is the hypothesis that this difference reflects a gap in the actual proficiency between the classes who were allocated to take V1 or V2 (but not both).

It had been expected that the average item scores would increase stepwise, given that items 1–2 were constructed from texts judged to be appropriate for beginner-level learners, items 3–4 for lower-intermediate learners, and items 5–7 for

upper-intermediate learners. However, while the scores for items 1–2 were indeed, on average, higher than those for items 5–7, the mean scores for the lower-intermediate items (3–4) were the highest in both V1 and V2.

It was concluded that these findings could represent a misjudgement of the difficulty levels of these texts and/or could indicate a practice effect occurring after the first two texts. To test the first of these hypotheses, online tools were used to analyse the complexity of the texts (Appendix D.2). Grotjahn (2002) recommends the type-token ratio (TTR; a measure of lexical diversity) as an indicator of complexity in C-tests. In addition, the researcher decided to consult a range of readability measures (Flesch Reading Ease Score, Gunning Fog Index, and Dale-Chall Score), which have also been used in C-test development (Sarapuu & Alas, 2016).

It must be noted that the accuracy of all of these measures increases with the length of the text, and that all items in this study were comparatively short (between 91 and 136 words). Nevertheless, the measures indicated a general trend of increasing complexity across the texts in both versions. Additionally, in both tests, the lower-intermediate texts (3–4) seemed more similar in complexity to the beginner-level texts (1–2), with a sharper increase in difficulty occurring with the upper-intermediate items (5–6). This supports the interpretation that the relatively high participant scores on items 3–4 in both versions arose due to a combination of the relative simplicity of these texts and a practice effect occurring after the first items.

7.1.2 Normality and Outliers

Before investigating the reliability and validity of the C-tests, the data were tested for normality and outlying values. This was a necessary first step, since the most commonly used extraction method in the planned Exploratory Factor Analysis

(Maximum Likelihood Estimation) assumes normality in the data (see e.g. Blunch, 2013; Fabrigar, Wegener, MacCallum, & Strahane, 1999). There are also robust estimation methods for non-normally distributed data, although outliers should still be treated (Finney & DiStefano, 2008).

The checks were carried out in SPSS separately for each version of the C-test. Normality was tested by inspecting the z-scores of skewness and kurtosis (see Appendix D.1). This procedure avoids the subjectivity of visual inspection by providing clear cut-off points, and is less vulnerable to sample size bias than other numerical methods, such as the Kolmogorov–Smirnov test (Field, 2018). In larger samples, such as this, z-scores above $|2.58|$ are commonly taken to indicate non-normality. Most items showed a high negative skewness and kurtosis, meaning that a larger proportion of the participants' scores clustered at the upper end of the distribution. The item scores in V2 were more skewed and peaked than those in V1 (i.e. clustered more closely together). There were four items with values beyond the cut-off point (printed in bold in Table D.1, Appendix D.1): one item in V1 and three in V2.

A common criterion for outliers is that they are more than 3.29 standard deviations removed from the sample mean (Field, 2018). Thus, the outliers in the current data were identified by transforming all item scores to z-scores in SPSS. Only two outlying scores were revealed, which were at the lower ends of the distributions of Text 3 and 4 in V2. To increase the accuracy of the subsequent analyses, the outliers were replaced with the cut-off values in each case, that is the cut-off scores, equivalent $z = -3.29$.

7.1.3 Validity: Exploratory Factor Analysis

As a measure of the two tests' initial validity, Exploratory Factor Analyses (EFA) were conducted separately for each version. The following criteria were used for judging the quality of multi-item scales through EFA (Hinkin, 1998; Worthington & Whittaker, 2006):

1. Items measuring the same factor should show moderate to strong intercorrelations ($r > .40$).
2. The proportion of individual item variances explained by the common factor should be at least 60% (communalities above .60).
3. Individual items should show clear factor loadings (above .40) and cross-loadings at least .15 below the main factor loading.

Principal Axis Factoring was chosen as the extraction method, because it is considered a 'robust' method, which can be used with non-normally distributed data (Finney & DiStefano, 2008). The anti-image correlation matrix and the Kaiser-Meyer-Olkin measure verified that the sample sizes were adequate for conducting these analyses ($KMO_{V1} = .91$ and $KMO_{V2} = .89$, respectively 'marvellous' and 'meritorious' according to criteria set out by Kaiser & Rice [1974]; all individual $KMO > .70$, which is above the commonly accepted minimum of .05 [Field, 2018]).

Next, the EFA was run to obtain eigenvalues for the factors underlying the data. In both versions, only one factor with an eigenvalue over Kaiser's criterion of 1 emerged (Kaiser, 1970). The emergent factor explained 66.93% of the total variance in V1 and 63.87% in V2, indicating that it was well represented in the items (Hinkin, 1998; Worthington & Whittaker, 2006). The scree plots also supported the decision to extract just one factor in each analysis.

It seems likely that the common factor represented the general language proficiency in both cases. An investigation of the extracted communalities (Table 7.1), however, showed that the proportion of individual item variances explained by the common factor was below the commonly recommended threshold of .60 (ibid.) in several cases. This indicated that the validity of the C-tests might be improved by excluding some texts from each version (this procedure will be described in Section 7.1.5 below). Nevertheless, the factor loadings in both versions were well above the minimal recommended threshold (above .40).

Table 7.1: Communalities and factor loadings in Exploratory Factor Analysis, C-tests Version 1 and 2. Extraction method: Principal Axis Factoring

<i>Item</i>	<i>Initial Communalities</i>	<i>Communalities after Extraction</i>	<i>Factor loadings</i>
Version 1 (V1)			
<i>Text 1</i>	.50	.51	.71
<i>Text 2</i>	.61	.59	.77
<i>Text 3</i>	.53	.52	.72
<i>Text 4</i>	.50	.48	.69
<i>Text 5</i>	.66	.69	.83
<i>Text 6</i>	.68	.74	.86
<i>Text 7</i>	.74	.80	.89
Version 2 (V2)			
<i>Text 1</i>	.52	.52	.72
<i>Text 2</i>	.59	.59	.77
<i>Text 3</i>	.54	.58	.76
<i>Text 4</i>	.41	.42	.66
<i>Text 5</i>	.67	.71	.84
<i>Text 6</i>	.70	.71	.85
<i>Text 7</i>	.53	.52	.72

7.1.4 Reliability: Cronbach's Alpha

A reliability analysis revealed that both tests had very good internal consistency (Cronbach's $\alpha_{V1} = .92$, $\alpha_{V2} = .90$). The item-total correlations were also calculated, as well as the new Cronbach's Alpha if any one of the items were to be deleted (Table 7.2). The right-hand column shows that the internal consistency of neither version would be improved by excluding any single item. However, as item reduction seemed necessary to improve validity, these numbers confirmed that the reliability would not be significantly compromised by excluding items with lower item-total correlations and factor loadings (e.g. Text 4 in each version).

Table 7.2: Item-total and reliability statistics, C-tests Version 1 and 2

<i>Item</i>	<i>Corrected Item-Total Correlation</i>	<i>Cronbach's Alpha if Item Deleted</i>
Version 1 (V1)		
<i>Text 1</i>	.69	.91
<i>Text 2</i>	.74	.90
<i>Text 3</i>	.69	.91
<i>Text 4</i>	.66	.91
<i>Text 5</i>	.79	.90
<i>Text 6</i>	.82	.90
<i>Text 7</i>	.85	.90
Version 2 (V2)		
<i>Text 1</i>	.67	.89
<i>Text 2</i>	.71	.88
<i>Text 3</i>	.71	.89
<i>Text 4</i>	.64	.89
<i>Text 5</i>	.79	.87
<i>Text 6</i>	.81	.87
<i>Text 7</i>	.68	.89

7.1.5 Item Exclusions

Based on the above evaluations of the C-tests, it was decided that the possibility of excluding two items from each version should be explored, to improve the validity of the two tests and make adjustments for the slight, but statistically significant, difference in the scores they produced. Upon re-examination of the item statistics (means, variance explained by the latent factor, item-total correlations, and factor loadings), it was decided that texts 1–4 were potential candidates for exclusion. To determine the effect that the exclusions would have on the validity and reliability of the tests, the EFA and Cronbach's alpha analysis were re-run six times, each time excluding two of the aforementioned four texts in different combinations. The results of these analyses are summarised in Appendix D.3.1.

The results for the versions that were finally chosen are reported in detail below. It was concluded that the best balance between overall test difficulty, validity, and reliability would be reached by excluding Texts 3 and 4 from both versions. A new variable was computed, which represented the overall percentile scores of the participants in this phase of the study without these items, and the tests of equality of means, validity, and reliability were re-run on both new versions of the test (henceforth Version 1.1. [V1.1.] and 2.1. [V2.1.]

The reliability analysis indicated that both tests had very good internal consistency (Cronbach's $\alpha_{V1.1} = .91$, Cronbach's $\alpha_{V2.1} = .88$). A separate EFA was also conducted on the five remaining items in each new version. The anti-image correlation matrix and the KMO once more verified the sampling adequacy for the analysis ($KMO_{V1.1} = .88$ and $KMO_{V2.1} = .81$, all individual item $KMO > .8$). Again, only one factor with an eigenvalue over 1 emerged in each version. The emergent factor now explained 73.2% of the total variance in V1, and 68.4% in V2.1. In both

versions, these factors also accounted for more than 60% of the variance in each individual item, indicating that the common factors were well represented in the data. Overall, the EFA indicated a good level of validity of the two modified C-tests, despite the comparatively small sample sizes. The new mean scores for V1.1 and V2.1 were much closer than before and a mixed model analysis revealed that the difference between them was no longer statistically significant ($F[1,58] = 3.60, p = .07$).

7.1.6 Adding More Difficult Items

As described previously (Section 6.2.1), the time participants were given to complete the C-test was reduced during the piloting phase because of the surprisingly large proportion of high-scoring students, from 35 minutes (five minutes per item) to 20 minutes (c. three minutes per item). Even at 20 minutes, however, 15% (8/54) of the total scores were still above 85%. This seemed problematic, especially given that the tests were intended to serve as repeated measures: For those participants already scoring at the high end on the first test, there would be little room for improvement.

Therefore, the researcher decided to add two more difficult ('advanced') texts to each version. Again, these texts were chosen with the help of the teachers from Phase I. Since the website that was the source of the other texts (www.onestopenglish.com) did not contain any advanced-level readings, the new items were based on excerpts taken from a different website (www.linguapress.com), which provided comparable materials. The same procedure as before was followed for turning these texts into items for the C-tests (see Section 6.1.1). Unfortunately, it was not possible to pilot these items in the same way before Phase III (due to access issues and time limitations on the project). However, adding a band of more difficult items is not uncommon in test development and re-running the above analyses with

the larger data set in Phase III confirmed the validity and reliability of these new tests (Sections 9.1.3.1 and 9.1.4).

7.2 Motivation and Attitudes Questionnaire

At this stage, the analyses of the motivation and attitudes data focused on (a) confirming the validity of the target constructs through EFA; (b) identifying which items best represented each of the constructs; and (c) subsequently excluding less well-fitting items to reduce the number of questions in the final questionnaire, thus minimising the burden on the participants.

7.2.1.1 Language Learning Orientations

Prior to running a principal axis EFA, the 34 language learning orientation items were inspected for normality and outliers. For nineteen items, the scores were non-normally distributed ($z_{skewness}$ and $z_{kurtosis} < |2.58|$), with most of them showing high skewness, that is, floor or ceiling effects. Through the inspection of z-scores, a total of seven outlying values, spread across six variables, were identified. They were replaced with the cut-off values, equivalent to $z = \pm 3.29$. At this stage, ten items were excluded from further analysis because they did not fit well with the overall item pool: They correlated poorly with the other items (i.e. fewer than ten correlations above .3) and shared only a small amount of the common variance (less than 25%).

Subsequently, a principal axis EFA was conducted on the 24 remaining LLO items. Based on the way that these orientations are theorised (as sub-components of language learning motivation), it was postulated that the underlying factors would be correlated. Hence, an oblique rotation (Direct Oblimin) was applied. The anti-image correlation matrix and Kaiser-Meyer-Olkin measure verified the sampling adequacy

for the analysis ($KMO = .91$; all individual item $KMO > .80$). The initial analysis identified four factors with eigenvalues above 1, which together explained 60.74% of the variance. As suspected, these factors were indeed moderately correlated ($r = [.28, .47]$).

However, there was a fifth factor with an eigenvalue of .99 (just under Kaiser's criterion of 1 [Kaiser, 1970]), which it seemed reasonable to consider for inclusion in the model. Upon closer inspection, it was revealed that the communalities (i.e. the amount of variance in each item explained by the extracted factors) were substantially higher in the five-factor solution, making this the preferred model (Table D.4, Appendix D.3.2.). The scree plot also supported the decision to extract these five factors. Together, they explained 63.98% of the total variance.

The factor loadings for the five-factor solution after rotation are reported in Table D.5, Appendix D.3.2. The items that clustered together suggest that the factors represent the following language learning orientations: (1) *instrumental obligation*; (2) *international/travel orientation*; (3) *instrumental benefit*; (4) *knowledge orientation*; (5) *intrinsic motivation*. The fact that five factors emerged, instead of the six which had been expected, is explained by the fact that both travel orientation and international orientation items loaded onto a single factor. This can be explained by the similar wording of these items and the previous qualitative finding that, most of the students in this population only had direct contact with people from international backgrounds when they were traveling abroad (see Section 5.1).

Not all items loaded onto the constructs they had been designed for. Usually, explanations could be found in the wording of these items: For example, Item 2 (*'I don't really like English'*) loaded onto instrumental orientation rather than intrinsic

motivation, perhaps because the participants interpreted the word 'English' as referring to the school subject, rather than the language (students also seemed to do this during the focus group interviews in Phase I).

Similarly, Item 15 (*'I want to know English because, in the future, I would like to live abroad [e.g. studying or working]'*) loaded onto instrumental benefits rather than international orientation, probably because the parenthetical addendum (*'e.g. studying or working'*) led participants to associate living abroad with having a successful career or receiving a good education. It was decided that this item should be reworded (i.e. the addendum deleted) and to include it in the final questionnaire, expecting that it would load onto international/travel orientation as planned. The loading of this new item was confirmed as part of the Confirmatory Factor Analysis in Phase III (Section 9.1.3.2).

At this point, some items were discarded to reduce the number of items in the final questionnaire, and ultimately the testing burden on the participants. Mostly, the items with the lowest factor loadings were excluded until three or four items (wherever possible) remained per factor (it is generally recommended that multi-item scales should comprise a minimum of three items, e.g. Dörnyei, 2009a). All of these subscales had good internal reliability (Cronbach's $\alpha > .8$) and it was found that the internal consistency of these scales would not be improved further by excluding any one item.

7.2.1.2 Motivational Selves

The participants' future and ideal L2 self were measured via ten questionnaire items, each of which the participants answered on two scales (measuring future goals and current self-perception). The resulting 20 'response-items' were first inspected for

normality and outliers. All of the items measuring the future L2 self, but none of the current self items, showed a significant positive skewness and kurtosis ($z_{skewness}$ and $z_{kurtosis} > |2.58|$), meaning that most of the participants' scores clustered at the lower end of the distribution. Unsurprisingly then, the inspection of z-scores revealed two to four outliers at the upper end of the distribution of each future self item (26 outliers in total), which were replaced with the values equivalent to $z = 3.29$. Further examinations, of the inter-item correlations and shared variance, did not identify any item which did not fit well within the overall pool. Therefore, all 20 response-items were included in the principal axis EFA. As with the previous analysis, a Direct Oblimin rotation was applied. The anti-image correlation matrix and the Kaiser-Meyer-Olkin measure verified the sampling adequacy for the analysis ($KMO = .89$; all individual item $KMO > .8$).

As expected, the analysis identified two latent factors with eigenvalues above 1, which together explained 72.30% of the variance. The scree plot supported the decision to extract these two factors. Table D.6, Appendix D.3.3 shows the factor loadings after rotation. The items that clustered together suggest that Factor 1 represents the current L2 self and Factor 2 represents the future L2 self. The correlation between these two factors was small ($r = .20$).

At this point, some items were discarded to reduce the length of the questionnaire and thus the testing burden on the participants. The questionnaire-items with the lowest factor loadings were excluded first (Item 1, '*I speak English fluently*' and Item 9, '*Speaking English is part of who I am*'). Notably, both of these items made specific reference to speaking skills, which left only two other items in the pool which addressed particular language skills (Item 4, '*I can write effortlessly in English*' and Item 7, '*I can read fluently in English*'). These items had been added to the initial

questionnaire only to counterbalance the bias towards speaking skills in the two other items, which were adapted from pre-existing questionnaires. It was therefore decided to remove the items pertaining to reading and writing from the pool as well.

After these exclusions, there were six remaining questionnaire-items to measure the current and ideal future L2 selves, with each subscale consisting of six response-items. These subscales had high internal reliability (Cronbach's $\alpha_{current} = .94$ and $\alpha_{future} = .96$). The item-total correlations and Cronbach's Alpha if any one item were to be deleted suggested that the internal consistency of these scales would not be improved further by excluding any one item.

7.2.1.3 Attitudes Towards Language Learning

Once again, checks for normality and outliers were first conducted with the items measuring attitudes towards language learning. Four of the informal learning and five of the formal learning items were non-normally distributed ($z_{skewness}$ and $z_{kurtosis} > |2.58|$). Most of these showed a clustering of scores at the upper end of the distribution. Nine outliers were also identified at the lower end of the distributions for three items measuring attitudes towards informal learning, which were replaced with the values equivalent to $z = 3.29$. Next, nine items were excluded from further analysis because they correlated poorly with the other items (i.e. fewer than five correlations above 0.3) and because they shared only a small amount of common variance (less than 20%).

Subsequently, a principal axis EFA with Direct Oblimin rotation was conducted on the remaining 12 items. The anti-image correlation matrix and the Kaiser-Meyer-Olkin measure verified the sampling adequacy for the analysis ($KMO = .85$; all individual item $KMO > .75$). The analysis identified three latent factors with

eigenvalues at or above 1 (Kaiser, 1970), which together explained 54.22% of the variance. However, the scree plot showed inflexions which seemed to support the extraction of two rather than three factors. For this reason, because the eigenvalue of the third factor was comparably small (1.00, i.e. right at the threshold of Kaiser's [1970] criterion), and because only one item seemed to load strongly and unambiguously onto this factor, it was decided to re-run the analysis extracting only two factors. Together, these explained 49.01% of the variance in the data.

Table D.7, Appendix D.3.4 shows the factor loadings for the two-factor solution after rotation. The item clustering together suggest that Factor 1 represents attitudes towards informal language learning, and Factor 2 attitudes towards formal language learning. A correlation, albeit small ($r = .20$) was found between these two factors. This can be attributed in part to Item 4 (*'Most of my English knowledge comes from school classes'*) and Item 5 (*'I find the English that we learn in school more useful than what I encounter in my free time'*), both of which showed strong main loadings ($> .5$) onto Factor 1, but also loaded noticeably onto Factor 2 ($> .35$). Perhaps these cross-loadings are not surprising, given their wording (particularly that of Item 5) and the fact that the two factors seem to be difficult to fully disentangle conceptually. However, the fact not all items whose wording included references to both formal and informal learning displayed such strong cross-loadings (i.e. Item 3: *'The English I encounter in my free time is more useful to me than what we learn in school'* and Item 6 *'I find the English I use outside of school is less important than what we learn in school'*) lends support to their theoretical distinction.

Items 4 and 5 were excluded from the final questionnaire because their strong cross-loadings meant that they were not suitable to distinguishing between attitudes towards informal and formal learning. After these final exclusions, five items

remained to measure each of the two factors (ten items in total). Both scales had satisfactory internal reliability (both Cronbach's $\alpha = .78$). The item-total correlations and the Cronbach's Alpha if any one item were to be deleted showed that the internal consistency of these scales would not be improved further by excluding any one item.

7.3 Informal Language Learning Diary

At this stage, the analyses of the informal language learning diary data focused mainly on the affective, cognitive, and linguistic engagement data. The goal was to develop multi-item scales that could be used to derive reasonably reliable and valid scores (scores averaged per participants, per week) for each of these aspects of engagement in ISLPs. The pilot version of the informal language learning diary had included twelve items (four per engagement construct) which the researcher intended to reduce to nine (three per construct) to lessen the burden on the participants in Phase III.

First the per-person, per-week averages for each of the twelve items were created and inspected for normality and outliers. All were approximately normally distributed ($z_{skewness}$ and $z_{kurtosis} < |2.58|$). Two outlying values were identified at the lower end of the distribution of the fourth affect and first cognition item, which were replaced with the value equivalent to $z = -3.29$.

A preliminary examination of the inter-item correlations and shared variances revealed one cognitive engagement item (*'It was [very easy – very difficult] to keep my mind on the activity'*) that did not fit well with rest of the items (i.e. few correlations above 0.3; lowest shared variance of all items at 38%). This item was therefore excluded from further analysis.

Subsequently, a principal axis EFA was conducted on the remaining eleven engagement items. Based on the way that these constructs are theorised, as different

aspects of engagement, it was postulated that the underlying factors would be correlated. Hence, an oblique Direct Oblimin rotation was applied. The KMO verified the sampling adequacy for the analysis ($KMO = .80$, all individual item $KMO > .60$). The EFA identified three factors with eigenvalues above 1, which together explained 62.40% of the variance. The scree plot supported the decision to extract these three factors.

Table D.8, Appendix D.3.5. shows the factor loadings after rotation. The items that clustered together suggest that Factor 1 represents *cognitive engagement*, Factor 2 represents *linguistic engagement*, and Factor 3 represents *affective engagement*. The factor correlation matrix confirmed that some of the extracted factors were indeed correlated: linguistic engagement only showed small correlations with the other factors ($r_{ling,cog} = .26$; $r_{ling,aff} = -.03$). Cognitive and affective engagement, however, were moderately correlated ($r_{cog,aff} = .39$). This can be explained, in part, by the strong cross-loadings of several items between these factors: While eight of the eleven items clearly loaded onto only one of the three emerging factors, three items showed medium loadings across two or three factors. This (in addition to the correlations between the factors themselves) confirms that there was some conceptual overlap between the three engagement dimensions. As such, the cross-loading items did not serve as good measures of any one type of engagement and were therefore discarded. Of the remaining eight items, two did not load onto the extracted factors as expected. These cases are briefly discussed in turn below:

1. Item 3 (*'The content was [not at all interesting – very interesting] to me'*): This item had been designed to measure affective engagement but loaded only moderately onto this factor. Instead, it loaded most strongly onto cognitive engagement. This could be due to the wording: The item specifically

referenced the '*content*' being interesting, which to the participants might have implied a degree of analytical thinking (i.e. cognitive engagement). This interpretation is supported by the fact that the word '*content*' appears in only one other item (*'All in all, I was [not at all — completely] focused on the content'*), which was designed and emerged from the analysis as a cognitive engagement item. It was decided that Item 3 would be included in the main study but reworded slightly (replacing the word '*content*' with '*activity*'), expecting that it might load onto affective engagement instead, which would then be represented by three items.

2. Item 4 (*'I was [not at all – completely] focused on trying to understand every single word'*): This had been designed as a linguistic engagement item but loaded most strongly on cognitive engagement in the EFA (item loading = .84), with merely a small cross-loading (.18) on linguistic engagement. This is perhaps unsurprising since there is some conceptual overlap between these two engagement constructs: Cognitive engagement is conceptualised here as capturing the strength of cognitive focus, while linguistic engagement is thought to relate more to the direction of this focus (i.e. focusing on language/form). The strong loading of this item onto cognitive engagement therefore seems to indicate that the participants focused on the phrase '*trying to understand*' (that is, strategic thinking), rather than on thinking about language in particular.

While these alternative interpretations offer an explanation for the cross-loading, the researcher decided that this was not an issue in the present case: It has been suggested that cross-loadings such as this are acceptable, as long as the difference between the stronger and weaker loading is at least .15

(Hinkin, 1998; Worthington & Whittaker, 2006). Furthermore, cognitive and linguistic engagement did emerge as separate factors from the analysis, which were only weakly correlated ($r = .26$) and shared no other substantial item cross-loadings.

After discarding the items that did not clearly load onto one factor, there were four items left to represent cognitive engagement but only two items to represent linguistic and affective engagement respectively. While these subscales showed acceptable internal reliability (Cronbach's $\alpha_{cog} = .85$; $\alpha_{ling} = .73$; and $\alpha_{aff} = .84^{13}$), the fact that affective and linguistic engagement were represented by only two items each could be perceived as problematic, given that it is generally recommended that multi-item scales should be comprised of a minimum of three items (e.g. Dörnyei, 2009a).¹⁴

However, as mentioned before, it was thought that the re-worded Item 3 would load onto affective instead of cognitive engagement, so that this factor would comprise 3 items in the final questionnaire. Furthermore, a new item was added to try and measure linguistic engagement: *'During this activity, I focused on [the general meaning – individual words/phrases]'*. The actual factor loadings of these modified/new items were examined as part of the preliminary analyses in Phase III (Section 9.1.3.5).

¹³ Generally, Cronbach's Alpha is not recommended for two-item scales, as it 'almost always underestimates the true reliability, sometimes rather substantially' (Eisinga, te Grotenhuis, & Pelzer, 2013). However, since the statistic returned acceptable values for both the linguistic and affective engagement scales in this case, the researcher did not think this was a problem in the present case.

¹⁴ Yet, it has also been shown on occasion that single- or dual-item measures can be reliable and valid if they are clear and unambiguous to the respondents (e.g. Ainley & Patrick, 2006).

7.3.1 Participant Compliance

In addition to the EFA and reliability analyses, the pilot data from the online learning diaries was analysed with regards to the total number of diary entries that each participant had submitted. This was to find out whether there was good representation of the participants in this part of the study, which relied on their independent participation at home. The number of entries per participant indicated that there were indeed some issues with compliance. Approximately one third of all participants in this phase (32/106) did not submit any diary entries at all. A further 44.4% (47/106) submitted six entries or fewer, which is less than one entry per day on average. The remaining 27.4% (29/106) submitted between seven and 34 entries per person. Of course, the simplest explanation for the large number of 'missing entries' is that participants may not have wanted to participate in this part of the study, or simply forgot. Several other potential reasons will be discussed in turn below, as well as the measures implemented to try to increase compliance in Phase III.

As noted by Bolger et al. (2003, p. 591), the major potential weakness of event-contingent diary designs is that participants may not reliably recognise each relevant event as such, and thus neglect to record some information. The written instructions for the participants in Phase II stated that they should 'make an entry every time [they came] into contact with English outside of school or their homework (e.g. watching a film or TV programme, playing a video game, listening to music, etc.)'. The examples provided in the instructions did not cover the full range of activities elicited in the informal language learning diary, which also covered practices such as using English social media, reading a book or an online text, hearing an English conversation, and speaking or writing in English. Thus, it was possible that some participants had difficulties identifying all of the activities they should record. The instructions were

therefore modified before Phase III to include a full list of the activities elicited in the online survey (Appendix B.3.2). Additionally, more extensive oral instructions were given during Phase III, wherein it was explained that the participants should also record any other English-language activities which the researcher might not have thought to put in the questionnaire.

A second potential problem with event-contingent diaries is that the participants may be tempted to generalise their experiences and provide fewer, aggregated reports, particularly if the relevant events occur frequently. This did indeed happen in the pilot study, where one participant made several entries listing all of the informal L2 practices with which they had engaged on a particular day and provided what were presumably estimated aggregate values for the duration and the cognitive, affective, and linguistic engagement items. To discourage future participants from doing the same, the written instructions were modified to emphasise that pupils should fill in the survey '**every time**' (emphasis in original) they were in contact with English outside of school.

Finally, it is possible that some participants did not make any diary entries, or submitted overall fewer entries, because they did not engage in any informal L2 practices on a particular day. Unfortunately, these cases could not be distinguished from those where participants may have forgotten or purposefully neglected to make an entry. To provide a way to distinguish these cases in Phase III, the participants were instructed to log in at the end of the day and choose '*I didn't use English outside of school today*' if they hadn't engaged in any of the activities that were of interest in the study. Finally, in the hope of increasing participation among pupils who may forget the diary, or who have low motivation to participate, the students in Phase III

were sent daily WhatsApp reminders for the duration of the diary data collection period (see Section 8.2).

8. Methods Phase III: Large-scale Quantitative Data Collection

8.1 Participants

For this phase of the study, data was collected from 354 pupils. They were recruited from 15 year 10 classes in six *Gymnasien* in rural districts in Lower Saxony. The participants' average age was 15.7 years ($SD = .66$). Among the 337 participants who provided information about their gender were 165 girls (49.0%), 171 boys (50.7%), and one participant who identified as non-binary. Therein, this sample was more well-balanced in terms of gender than the groups in Phase I and Phase II. In terms of their language background, of the 336 participants for whom this data was available, 298 (88.7%) had been raised monolingually in German. Of the rest, five (1.41% of the sample) grew up bilingually in German and English and 29 (8.63%) speaking German and one other language (most commonly Russian or Kurdish). There were also four participants who were not native speakers of German: Three recent immigrants, who had moved to Germany and started learning German within 3 years before the start of this study, and one Italian foreign exchange student, who had been studying German for eight years.

All of the participants knew at least one other language in addition to German and English,¹⁵ which was expected given that all students at a *Gymnasium* need to study at least two foreign languages (usually English and either Latin, French, Spanish, or Italian). While 278 participants (82.7%) knew only the three languages required for students at this level, a significant number (58 participants, 17.3%)

¹⁵ A 'known language' was any language for which the participants reported a self-assigned proficiency score of at least two out of seven (the end points of the scale were marked '*I only know a few phrases*' and '*native speaker-like*'). Languages in which the participants indicated a proficiency score of one were excluded because many listed a large number of such languages, which could be indicative of a response bias (i.e. wanting to appear more linguistically talented).

indicated knowledge of four or more languages. Unsurprisingly, this group was made up mostly of bilinguals (31/58, 53.4%), but it also included a large number of monolingually raised participants (27, 46.6%) who had learned multiple second languages after the age of four. At the time of the study, most of the pupils had been learning English for seven or eight years ($M = 7.98, SD = 1.20$). As in the previous phases, these students had CEFR B1-B2 level proficiency, on average, according to their teachers' estimates.

8.2 Procedure

The researcher visited each of the participating classes for two lessons (45 minutes) at two time points. The lessons began with a brief introduction of the researcher and the study and answering the students' remaining questions, if any. At this point, a brief explanation of the informal language learning diary (see below) was also given, and the sheets containing the written instructions and individual participant IDs were distributed (Appendix B.3.1). Then, the participants received oral and written instructions for the C-tests (Appendix B.1), which they were given 20 minutes to complete. During this time, the pupils were reminded periodically to move on to another text if they felt stuck, and that they should not try to copy answers from one another. The two versions of the C-test were administered in counter-balanced order: eight out of the fifteen classes completed V1 at T1 and V2 at T2; the seven remaining classes were given V2 before V1. After turning in their test papers, the pupils had the remaining part of the lesson (approximately 20 minutes) to complete the questionnaire (Appendix B.2). They were encouraged to ask the researcher or their teacher if anything was unclear to them.

The participants were instructed to document their informal contact with English using the informal language learning diary over the seven days immediately following the first classroom visit. They were asked to submit an entry every time they came into contact with English outside of school, through activities such as using social media, watching a film or serials, listening to music, etc (see Appendix B.3.2). In the hope of increasing completion rates of this part of the study beyond those in Phase II (Section 6.1.3.) consenting participants were sent daily reminders on their phones:¹⁶ In most classes, the researcher collected the individual students' mobile phone numbers and sent them individual messages using WhatsApp. However, some classes who already had a WhatsApp group they used to communicate with each other and with their teachers instead decided to nominate one volunteer to receive the daily messages from the researcher and forward them to the whole class via the group chat. Approximately 10% of the participants did not receive any reminders as they or their parents did not consent to the sharing of mobile phone numbers for this study. Since the numbers were collected anonymously and completely independently of all other data (see Section 3.4), it is not possible to identify which particular pupils did not receive the reminders and whether they submitted fewer diary entries than the other participants.

8.3 Data Analysis

After the data collection was completed, the C-tests were marked on paper and the scores, together with the questionnaire responses, were entered into IBM SPSS Statistics. This data set was then merged with the data downloaded from Qualtrics,

¹⁶ The participants interviewed in Phase I had identified WhatsApp messages as their preferred method for receiving reminders, over e.g. email, which they reported they did not check regularly.

containing the participants' informal learning diary entries. Any necessary data transformations and the more basic statistical analyses were carried out in SPSS, while more advanced tests (CFAs and SEM) were run in R. Table 8.1 provides an overview of the analytical approaches and the purpose they served in the context of this study. The rationale for choosing these approaches will be discussed at the start of each of the respective results sections.

Table 8.1: Overview of analytical approaches

<i>Aspect of Study</i>	<i>Analytical Approach</i>	<i>Software</i>
Preliminary Analyses (Section 9.1)		
Missing data and exclusions; normality checks	Investigating the available data through descriptive statistics	SPSS
Investigating measurement quality: reliability and validity	Cronbach's Alpha; Confirmatory Factor Analysis	SPSS SPSS R (lavaan)
Research Question 1 (Section 9.2)		
<i>How do learners of English as a second language, attending year 10 classes at German Gymnasien, engage in English informal second language practices?</i>	Descriptive analyses, e.g. time spent on ISLPs; most popular activities	SPSS
<i>Can groups of learners be identified, who show distinct patterns of engagement in informal second language practices?</i>	Latent Profile Analysis	R (tidyLPA)
Research Question 2 (Section 9.3)		
<i>What are the direct relationships between language proficiency and engagement in informal second language practices?</i>	Structural Equation Modelling	R (lavaan)
Research Question 3 (Section 9.4)		
<i>What are the direct relationships between language learning motivation and attitudes towards language learning and engagement in informal second language practices?</i>	Structural Equation Modelling	R (lavaan)
Research Question 4 (Section 9.5)		
<i>How do language proficiency, language learning motivation, and attitudes interact in their relationships with engagement in informal second language practices?</i>	Structural Equation Modelling	R (lavaan)

9. Findings Phase III: Quantitative Analyses

This chapter presents the results the Phase III quantitative data analyses. Section 9.1 discusses the preliminary tests that were conducted to investigate the completeness of the data set, normality and outlier checks, Confirmatory Factor Analyses which served as tests of validity, and finally reliability analyses. Section 9.2. addresses the analyses carried out to address Research Question 1, mapping the patterns of engagement in informal second language practices among the students in this study. Sections 9.3–9.5 address Research Questions 2–4, exploring, step-by-step, the relationships between engagement in ISLPs, general L2 proficiency, language learning motivation, and attitudes towards informal and formal learning.

9.1 Preliminary Analyses

9.1.1 Completeness of Data

While a total of 354 students participated in this phase of the study, fewer full data sets were available as some pupils did not complete one or more stages of the data collection (i.e. the motivation and attitudes questionnaires and C-tests at time 1 [T1] or time 2 [T2], or the informal language learning diary). Table 9.1 shows the number of participants who took part in each stage. The attrition between T1 and T2 was explained by pupils being ill or otherwise absent from school. There are several possible reasons for the lower number of participants in the informal language learning diary (cf. Section 6.1.3), including the high demands that this data collection method placed on students in terms of time and self-directed effort. In addition to these missing data, some participants (twelve who were enrolled in several English-medium classes and five who were raised bilingually in German and English) were excluded from the analyses concerning language proficiency.

Table 9.1: Overview of the number of participants in each stage of Phase III

	Number of participants	Proportion of the total number
<i>Total number of participants across all stages</i>	354	
<i>Questionnaire & C-test time 1 (T1)</i>	335	94.6%
<i>Questionnaire & C-test time 2(T2)</i>	294	83.1%
<i>Informal Language Learning Diary*</i>	262	74.0%
<i>Participants who completed all three stages</i>	213	60.2%

*Participants who submitted diary entries on at least one day.

9.1.2 Normality and Outliers

Identifying outliers and non-normally distributed variables was a necessary first step, since the most commonly used extraction method in the planned CFA and SEM analyses (Maximum Likelihood Estimation) assumes normality in the data (Blunch, 2013; Fabrigar et al., 1999). There are also methods which are robust to non-normality, although outliers should still be treated (Finney & DiStefano, 2008). As in Phase II, outliers were identified by transforming the measurement variables to z-scores (cut-off: $z > |3.29|$). However, as the total sample size now exceeded 300, it was no longer appropriate to rely on the z-scores of skewness and kurtosis to identify non-normal distributions as before (Field, 2018). Instead, the absolute values of skewness and kurtosis were consulted. The normality of the C-test scores was inspected separately for both versions and time points (four groups). The motivation and attitudes variables also inspected separately for time 1 and time 2.

One to four outlying values were identified at the lower end of most of the variables. They did not always stem from the same participants. To increase the accuracy of the subsequent analyses, the outlying values were replaced with the cut-off values in each case, that is the scores equivalent to $z = -3.29$. In most cases, the

outliers were scores at the lower or upper end of the ten-point Likert-type scales where the majority of the other participants' answers clustered in a narrow range (i.e. the overall distribution was leptokurtic, see below).

The tables in Appendix E.1 list the descriptive statistics for all of the tested variables after treatment of the outliers. All C-test items and questions measuring attitudes towards language learning were approximately normally distributed. Among the language learning orientation items, the distribution of only one item (*'I don't need to know English well because I don't like traveling to other countries'*) was non-normal, showing a high positive skewness and kurtosis (floor effect). That is, most participants chose to 'completely disagree' with this statement. A similar pattern was found for the only non-normally distributed engagement item (*'All in all, this activity was [boring – entertaining]'*), where most responses clustered at the positive end of the scale (ceiling effect). Finally, most of the scales measuring the future L2 self showed significant positive kurtosis and negative skewness, indicating that a vast majority of the participants had a strong ideal future self, that is they wished to be highly proficient in English.

In sum, since a number of variables were found to be non-normally distributed, the researcher decided to employ robust Maximum Likelihood Estimation during the CFA and SEM analyses (Blunch, 2013; Fabrigar et al., 1999).

9.1.3 Validity: Confirmatory Factor Analysis and Measurement Invariance

Modelling

To strengthen the claim for the validity of the instruments developed in this study, Confirmatory Factor Analyses (CFA) were conducted at this stage: This technique statistically tests how well previously specified factor structures are represented in a

data set. The CFAs were run in R using the lavaan package and robust Full-Information Maximum Likelihood estimation (FIML) to impute missing data points. As is common practice in CFA, the factor loading of the first item measuring each latent construct was fixed to one to allow for model identification (e.g. Blunch, 2013; Byrne, 2010; R. B. Kline, 2015).

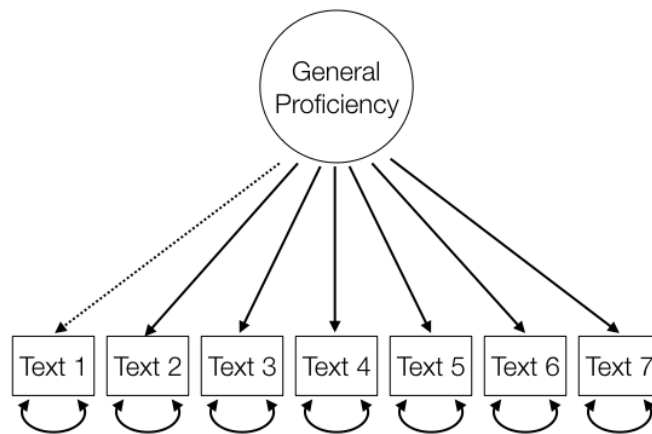
Furthermore, the school classes to which the participants belonged were included in these SEM models (and all subsequent SEM analyses) as a clustering variable: This is common in educational studies to control for the influence of group-level variations such as teacher/classroom practices and student group dynamics on the individual pupils' scores (Hedges & Rhoads, 2010). In particular, the researcher expected that there might be substantive differences between the classes in terms of their average language proficiency. The effects of such nested sample structures can be investigated in detail by employing more complex, multi-level models. However, since these group differences were not the focus of this study, specifying a clustered design (which prompts the model to adjust the fit indices and parameter estimates for the effects of between-group variance; Rosseel, 2018) was chosen as a simpler alternative. Post-hoc analyses of the *intra-class correlations* (Appendix E.2) confirmed that clustering in classrooms did indeed have a substantial effect on a number of the C-test items (Texts 2– 4 in V1, Text 5 in V2).

9.1.3.1 C-tests

Given that C-tests are thought to measure general language proficiency, and based on the fact that only one latent factor emerged in the EFA in Phase II (Section 7.1.3), the researcher specified the one factor model pictured in Figure 9.1. Subsequently, four

CFAs were conducted: one for each C-test version (V1 and V2) at each of the two time points (T1 and T2).

Figure 9.1: One-factor measurement model for C-tests



*The dashed arrow above Text 1 indicates that this factor loading was fixed to 1 to allow for model identification.

Table 9.2 contains an overview of the *goodness-of-fit indices* for these four analyses. A brief explanation of each of these indices is given below.

The *Chi-square (χ^2) Test* is a measure of the difference between the observed item covariances and those predicted by the model: A non-significant chi-square statistic ($p > .05$) is thus desirable, indicating that the expected covariances do not differ significantly from those observed. However, the chi-square test is extremely sensitive to sample size, wherefore alternative fit measures are often considered to be more reliable (Byrne, 2010). Among these are the *Comparative Fit Index* (CFI) and the *Tucker-Lewis Index* (TLI), both of which compare the researcher-specified model to a saturated model (wherein all components are related to all other components, resulting in maximum fit) and an independence model (wherein all variables are

unrelated, resulting in minimum fit). Generally, values above .90 are considered ‘acceptable’ and values above .95 ‘excellent’ (Blunch, 2013; Byrne, 2010).

Further important fit indices are the *Root Mean Square Error of Approximation* (RMSEA) and the *Standardized Root Mean Squared Residual* (SRMR). Both measure how closely the model reproduces the covariances among the observed variables. Lower values are desirable because they indicate that the differences between the observed and estimated covariances — the residuals — are small. Commonly cited critical values for these measures are .06 for a ‘good’ fit, or .10 for an ‘acceptable’ one (e.g. Blunch, 2013; Byrne, 2010; Hu & Bentler, 1999). However, it should be noted that the RMSEA is not reliable for models with very few degrees of freedom ($df < 10$), when its values tend to become inflated (Kenny, Kaniskan, & McCoach, 2015).

Table 9.2: Goodness-of-fit indices, CFA of C-tests V1 and V2 at T1 and T2

<i>Model</i>	<i>n</i>	χ^2	<i>df</i>	<i>p</i>	<i>CFI</i>	<i>TLI</i>	<i>RMSEA</i>	<i>SRMR</i>
VERSION 1, TIME 1 (V1T1)								
<i>Baseline (as planned)</i>	170	21.88	14	.081	.98	.98	.06	.03
<i>Correlated residuals Intrinsic3, Intrinsic4</i>	170	10.80	13	.628	1.00	1.00	.00	.02
VERSION 2, TIME 1 (V2T1)								
<i>One-factor model</i>	148	59.62	14	< .001	.86	.79	.16	.07
<i>Correlated residuals Intrinsic3, Intrinsic4</i>	148	14.35	13	.350	1.00	.99	.03	.04
VERSION 1, TIME 2 (V1T2)								
<i>Baseline (as planned)</i>	108	11.76	14	.626	1.00	1.00	.00	.03
VERSION 2, TIME 2 (V2T2)								
<i>Baseline (as planned)</i>	150	33.04	14	< .01	.97	.96	.08	.03

The CFA of the hypothesised one-factor model achieved excellent fit for all data sets except in the case of V2T1. This generally good fit and the relatively high individual item factor loadings (see Figures E.1–E.6, Appendix E.3) support the claim that one latent factor (presumably *general language proficiency*) could explain a substantive

portion of the variation in the item scores. In other words, the CFA suggests good structural validity in the case of all tests except V2T1. In that case, the aforementioned indices signified a bad fit of the one-factor model and some of the factor loadings were not as strong as in the other analyses, suggesting that the latent factor accounted for less variation in this data set than the others.

The Modification Indices (MI) supplied by the lavaan software were consulted to gain an understanding of why the model did not fit the V2T1 data well. There is some disagreement about the extent to which it is appropriate to make changes to CFA models based on these indices. Some argue that researchers should return to EFA to find, and subsequently re-test a better fitting model (e.g. Browne, 2001; MacCallum, Roznowski, & Necowitz, 1992). Others find it acceptable to exclude items identified as contributing to bad model fit, or to correlate some error variances within the same factor, but also warn that changes should be applied only where a theoretical explanation can be found (e.g. Blunch, 2013; Byrne, 2010; Schmitt, 2011). In this and the subsequent analyses, the researcher decided to consult the modification indices to identify possible sources of model misfit whenever necessary and to consider the suggested changes in light of the theoretical definitions of the latent constructs, the wording of the individual items, as well as the findings of the EFAs in Phase II.

In the case of the V2T1 data set, the software suggested correlating the error variances of Texts 6 and 7, implying that there was some covariance between these items not captured by the latent factor. It is worth noting that the MI also suggested correlating Texts 6 and 7 in the analysis for V1T1, but not for either of the T2 analyses. This implies that there might be a common factor besides *general language proficiency* which affects participants' performance on Texts 6 and 7 at T1.

Since these were the last two items on the test, it seems plausible that this might represent an effect of test-taking speed: The C-test format was completely new to the participants and the test was timed, so that pupils who had more difficulty getting used to the format might not have had time to complete Texts 6 and 7 to the best of their abilities, resulting in lower scores. The fact that these correlations significantly affected the model fit at T1, but not T2 supports this theory, since the participants were more familiar with the test format the second time around. When the model was re-specified to allow for residual correlations between Texts 6 and 7 at T1, the fit was significantly improved for both V1 and V2 ($\Delta\chi^2_{V1T1}[1] = 5.97$, $p < .05$; $\Delta\chi^2_{V2T1}[1] = 45.93$, $p < .001$; see Table 9.2 for complete fit indices).

The CFAs presented above show that the modelled latent factor was well represented in the individual C-test items. This strengthens the claim that the C-tests were valid measurement tools. However, there is a second major concern with establishing measurement validity, which is not addressed by the methods employed so far: whether or not the data collected across the two time points and with the two C-test versions is comparable. This can be tested using a form of *Multi-Group Confirmatory Factor Analysis* called *Measurement Invariance Modelling* (MIM) (e.g. Byrne, 2010; van de Schoot, Lugtig, & Hox, 2012; Vandenberg & Lance, 2000). MIM offers a way to assess whether measurement instruments function in the same way over time or across different groups of participants. In this study, MIM was employed to demonstrate that the participants' scores across the different C-test versions (which were counter-balanced) seemed to adhere to the same latent factor model.

To establish measurement invariance, increasingly restrictive multi-group models are fit to the data (see Appendix E.4). Statisticians generally require that a

good fit needs to be achieved at the most restrictive level (*strict invariance*) in order to conduct valid mean comparisons across groups or time, although practitioners agree that this can rarely be achieved in most Social Science research (van de Schoot et al., 2012). Nevertheless, both researchers and statisticians concede that achieving the less restrictive *metric invariance* is enough to justify conducting analyses based on variances and co-variances, such as the planned Structural Equation Modelling (Bialosiewicz, Murphy, & Berry, 2013).

To account for the counter-balanced longitudinal design, measurement invariance of the C-tests was tested in three ways (across groups, across versions, and across times, see Figure E.8, Appendix E.4). Each analysis achieved a good fit at the level of metric invariance, or beyond. This strongly suggests that the two C-test versions provided valid data for the planned SEM analyses. That is, it seems justifiable to apply the same factor model (introduced above) to the data collected with both test versions, from both sub-samples of participants, at both time points in order to estimate general language proficiency.

9.1.3.2 Language Learning Orientations

The questionnaire developed in Phase II included a total of 22 items designed to represent five language learning orientations: four items each for intrinsic motivation, instrumental obligation, instrumental benefits, and knowledge orientation and six items for measuring international orientation. To further test the validity of this instrument, two CFAs were conducted, separately fitting the same five-factor model to the data collected at T1 and T2 in Phase III. Table 9.3 (p. 179) reports the goodness-of-fit indices for these analyses.

A good fit was achieved for the data at both points after some modifications to the model. These were necessary because the initially hypothesised model did not, at first, achieve a good fit and the software used to conduct the analysis gave a warning that the model may be too complex to achieve accurate identification. Therefore, modification indices were consulted to gain an understanding of where the tested model diverged from the patterns in the collected data, with the aim of excluding some of the items that fit less well, improving the model fit and reducing complexity. Stepwise changes were implemented (see Table 9.3) where both the T1 and T2 model indicated the greatest misfit:

1. Obligation 4 (*'I don't really like English'*) was excluded, because of its lower factor loading onto instrumental obligation and heavy cross-loading onto intrinsic motivation. This replicated what was found in the EFA during Phase II (Section 7.2.1.1) and is likely due to the participants interpreting *'English'* either as referring to the language itself (implying intrinsic motivation) or the school subject (implying an obligation to learn). In any case, this cross-loading showed that the item did not accurately distinguish between different LLOs, so that the researcher chose to exclude it.
2. Correlations between the error variances of two items representing intrinsic motivation (Intrinsic 3: *'I find learning English interesting'* and Intrinsic 4: *'I think learning English is really fun and see it as a challenge'*) were added to the model, implying that these items measured a common concept beyond intrinsic motivation. As these two questions specifically referred to *'learning English'*, it could be argued that they measure intrinsic language *learning* motivation, while the other two intrinsic motivation items (*'I find English*

exciting and I don't really like the English language') focused more on affect related to the language itself.

3. International 6 (*'I don't need to know English well, because I don't plan on ever living abroad'*) was excluded due to lower factor loadings on international orientation and comparatively high cross-loadings onto the instrumental benefits factor. This also emerged in the Phase II EFA (Section 7.2.1.1), whereupon the researcher speculated that the cross-loading might be caused by participants associating living abroad with having a successful career (instrumental benefit). The item was included in the Phase III questionnaire mainly to confirm this cross-loading with a bigger sample but was now excluded from further analysis.
4. Similarly, International 3 (*'I don't need to know English well, because I don't like traveling to other countries'*) was excluded due to its very low loading onto the international orientation factor. This might be due to the substantial ceiling effect for this item: a vast proportion of the scores clustered at the upper bound of the ten-point scale ($M = 8.9, SD = 1.55, S = -2.05, K = 3.70$). It seems that most participants were exceedingly enthusiastic about traveling, meaning that this item was not good at distinguishing between individuals and should be discarded.
5. Knowledge 4 (*'I want to know English so that I can acquire new ideas and broaden my horizon.'*) was excluded for showing strong cross-loadings onto several other factors, primarily international orientation, but also intrinsic motivation and instrumental benefits. Again, this might be due to the item wording: *'Acquiring new ideas'* related, fairly literally, to knowledge orientation (wanting to know English to help one to become better educated

and pursue personal interest). Meanwhile, '*broadening [one's] horizon*' can be interpreted in various ways, for example referring (as intended) to increasing one's range of knowledge, but also to socialising with different people, such as by traveling to different countries. The item might have been a more accurate measure of knowledge orientation without the second statement, but as it was, the researcher felt it warranted exclusion.

6. Finally, Intrinsic 1 (*'I find English exciting'*) was excluded for showing a strong, inverse cross-loading onto instrumental obligation. Similar to the Obligation 4 item discussed above (pt. 1), this cross-loading can likely be attributed to participants interpreting '*English*' as either referring to the language or the school subject.

Altogether, these changes significantly improved the fit of both models ($\Delta\chi^2_{\tau_1}[85] = 360.80, p < .001$; $\Delta\chi^2_{\tau_2}[85] = 370.21, p < .001$), with all indices except for Chi-square statistic reaching the range indicative of an 'acceptable' to 'good' model fit. While making additional changes might have further improved the fit, the model was accepted at this point for two reasons: First, further item exclusions would have reduced the number of items per factor to fewer than the recommended minimum of three. Second, the modification indices for the two analyses diverged substantially at this point, each suggesting different adjustments, which were not as clearly justified through prior EFA findings or a closer look at the item wording, as above. The factor loadings, error variances, and inter-factor correlations of this final model, applied to both time points, are shown in Figures 9.2 and 9.3.

Table 9.3: Goodness-of-fit indices, CFA of language learning orientations at T1 and T2

Model	Time 1 (n = 326)							Time 2 (n = 281)						
	χ^2	df	p	CFI	TLI	RMSEA	SRMR	χ^2	df	p	CFI	TLI	RMSEA	SRMR
<i>Baseline (as planned)</i>	587.25	177	< .001	.87	.84	.09	.06	582.96	177	< .001	.86	.83	.10	.07
<i>Excl. Obligation 4</i>	458.91	158	< .001	.90	.88	.08	.06	472.39	158	< .001	.85	.82	.10	.07
<i>Correlated residuals Intrinsic3, Intrinsic4</i>	420.69	157	< .001	.91	.89	.08	.06	421.16	157	< .001	.90	.88	.08	.07
<i>Excl. International 6</i>	361.14	140	< .001	.92	.90	.07	.06	376.13	140	< .001	.90	.88	.08	.06
<i>Excl. International 3</i>	316.15	123	< .001	.93	.91	.07	.05	295.26	123	< .001	.92	.91	.07	.06
<i>Excl. Knowledge 3</i>	284.25	107	< .001	.93	.91	.07	.05	229.27	107	< .001	.94	.93	.07	.05
<i>Excl. Intrinsic 1</i>	229.45	92	< .001	.94	.92	.07	.05	212.75	92	< .001	.94	.93	.07	.05

Table 9.4: Goodness-of-fit indices, MIM of language learning orientations at T1 and T2

Model	χ^2	df	p	CFI	TLI	RMSEA	SRMR
<i>Configural Invariance</i>	696.95	399	< .001	.94	.93	.05	.06
<i>Metric Invariance</i>	729.70	410	< .001	.94	.92	.05	.07
<i>Scalar Invariance</i>	795.19	426	< .001	.93	.92	.05	.07
<i>Strict Invariance</i>	810.54	442	< .001	.93	.92	.05	.07

Finally, a series of MIMs was run to further investigate the similarity of the fit of this updated five-factor model across the two time points (Table 9.4). In addition to the latent factors, the residuals of equivalent items were correlated across the time points (e.g. Intrinsic 2.T1 with Intrinsic 2.T2). Although the χ^2 test indicated a significantly worse fit of the metric and scalar invariance model than their respective predecessors ($\Delta\chi^2_{metr,conf}(11) = 32.75, p < .001$; $\Delta\chi^2_{scal,metr}(16) = 65.49, p < .001$), the strict invariance model still achieved acceptable fit in terms of the other indicators. This supports the assumption that the indicators measure the latent factors in similar ways at both time points, so that comparisons can be made via analysis of variances and covariances (i.e. SEM).

Figure 9.2: Final five-factor model of language learning orientations at T1

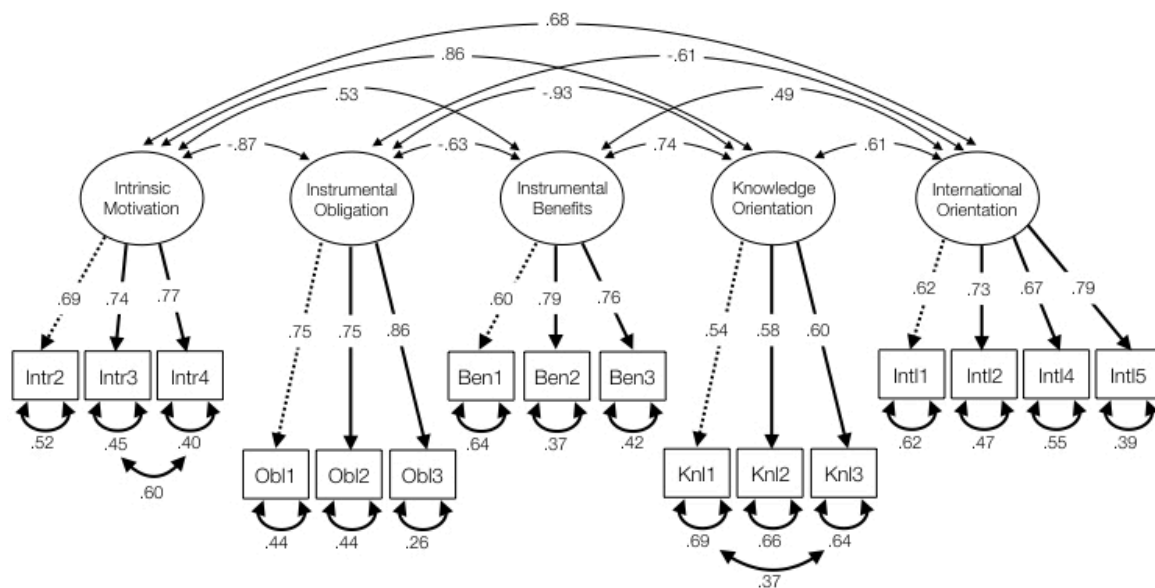
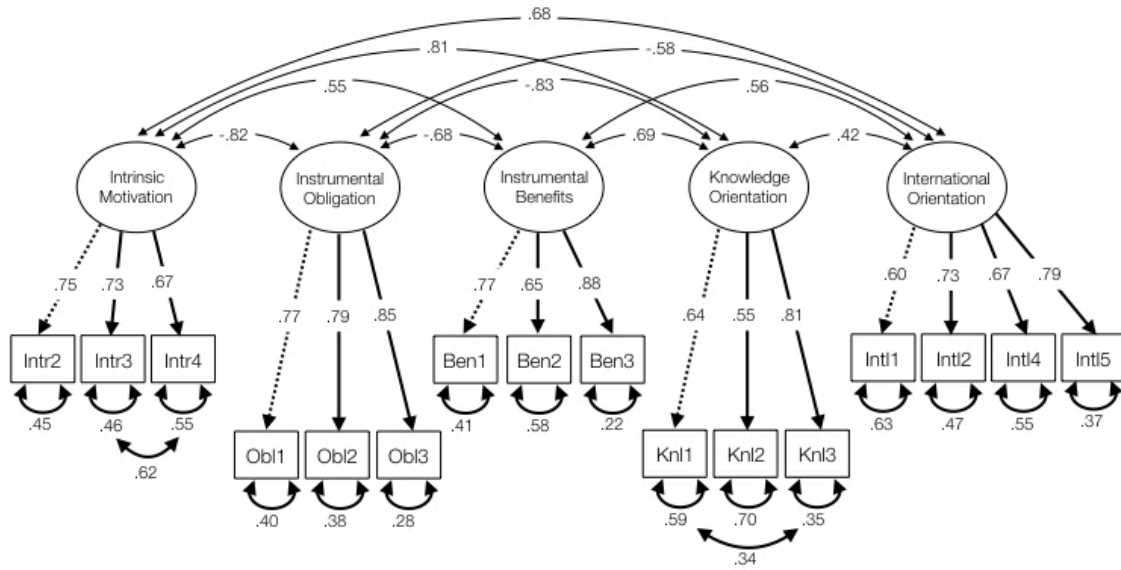


Figure 9.3: Final five-factor model of language learning orientations at T2



9.1.3.3 Second Language Motivational Selves

The participants' future and ideal L2 self were measured via six questions, each of which the participants answered on two scales (measuring future goals and current self-perception). The resulting twelve scores were entered into separate CFAs for the T1 and T2 data according to the two-factor model pictured below. To account for effects of item wording, the error variances of the current and future self scales that responded to the same question were allowed to correlate. At both T1 and T2, the planned model achieved an 'acceptable' fit on the TLI and RMSEA statistics, and a 'good' fit according to the CFI and SRMR (Table 9.5, p. 188). The factor loadings and residual variances, however, indicated that the second item in both the current and future ideal self inventories (*I can express my whole personality in English*) were predicted by the proposed latent factors substantively less well than the other items (factor loadings: .54 and .68 vs. .78 and above). As shown in Table 9.5, the exclusion of these items significantly improved the model fit in terms of the Chi-square statistic ($\Delta\chi^2_{T1}[18] = 51.57, p < .001$; $\Delta\chi^2_{T2}[18] = 73.60, p < .001$) and also improved (if

only slightly) some of the other fit scores. The factor loadings, residual variances and covariances resulting from this adapted model are shown in Figures 9.4 and 9.5.

Figure 9.4: Final two-factor model of L2 motivational selves at T1

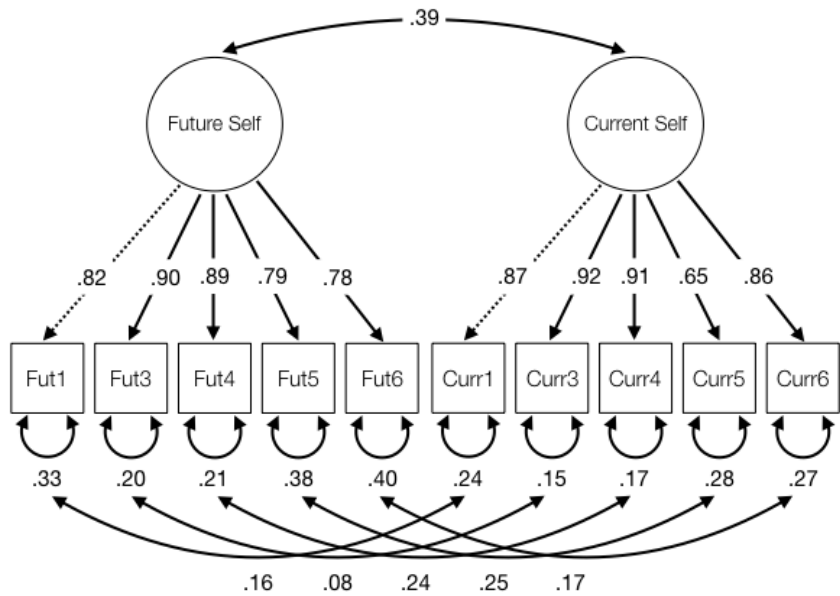
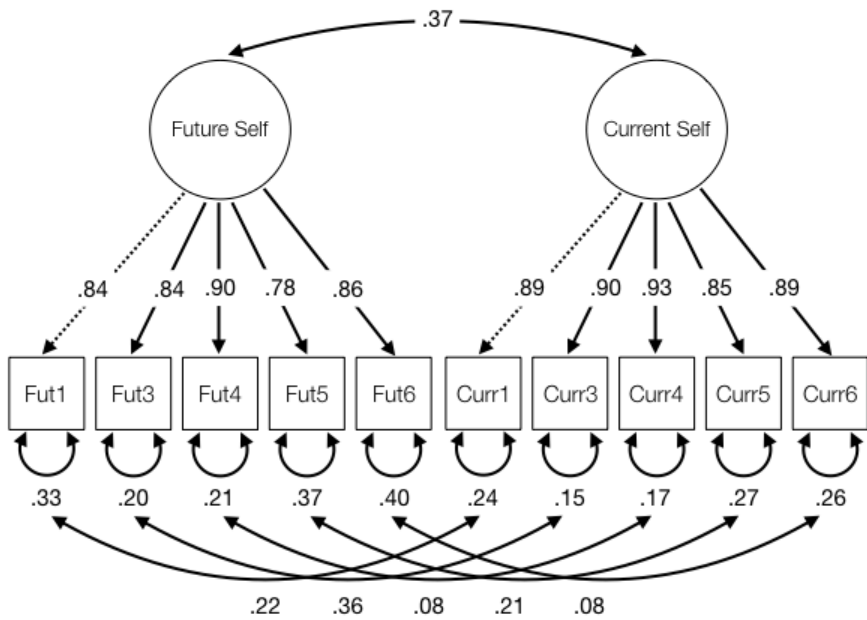


Figure 9.5: Final two-factor model of L2 motivational selves at T2



With this data, too, a series of longitudinal MIMs was tested to investigate the similarity of the model fit across the two time points (Table 9.6, p. 188). In addition to the residual correlations between the future and current self items within the same time point (pictured above) all indicators were also correlated with their counterparts across time (e.g. Future 1.T1 with Future 1.T2). The χ^2 test indicated that the fit of the scalar invariance model was significantly than that of the metric invariance model ($\Delta\chi^2_{scal,metr}[10] = 22.99, p < .05$). Nevertheless, even the most restrictive (strict invariance) model still achieved a good fit across all of the other indicators, supporting the equivalency of the questionnaire across the two time points.

Motivational distance

As discussed previously (Section 2.3.1), the possible selves theory on which Dörnyei's (2005) Motivational L2 Self-System is based hinges on the idea that learners' behaviour is motivated by a (mostly unconscious) desire to reduce the distance between the current and ideal future selves. Therefore, the current and ideal future self scales were designed with a view towards also deriving a measure of the *motivational distance*. This was done by subtracting the current self scores from the corresponding future self scores, resulting in six scores for the motivational distance. Their loading onto one common factor (Figures 9.6 and 9.7) was subsequently analysed via the same CFA methods as above.

Figure 9.6: Final one-factor model of L2 motivational self-distance at T1

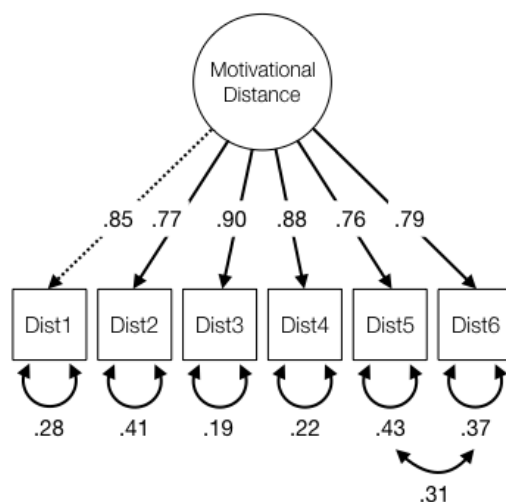
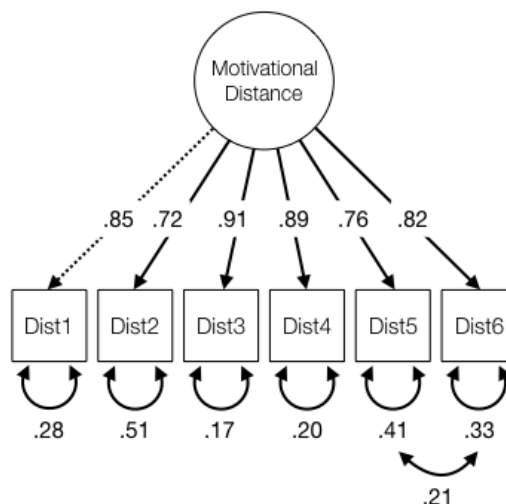


Figure 9.7: Final factor model of L2 motivational self-distance at T2



The model achieved an ‘excellent’ fit when applied to the T2 data and an ‘acceptable’ fit level at T1 (see Table 9.7, p. 188). The modification indices suggested that the T1 fit could be improved substantially by correlating the residuals of the last two items. With the inclusion of this change, an excellent fit was achieved at both T1 and T2.

Next, a series of MIMs was also applied to this data (see Table 9.8, p. 189), with the residuals of the corresponding items correlating across time points (e.g. Distance 1.T1 with Distance 1.T2). Although the χ^2 statistic indicated a significantly worse fit of the metric invariance model than the configural model ($\Delta\chi^2_{metr,conf}[5] = 11.72, p < .05$), even the strict invariance model still achieved a

good fit in terms of the other indicators. This supports the assumption that the six indicators measure the latent factor in a sufficiently similar manner at both time points to allow for comparisons via analyses of variances and covariances.

9.1.3.4 Attitudes Towards Language Learning

There were a total of ten questionnaire items which measured participants' attitudes towards language learning: five focused on formal and five on informal learning. Once again, two CFAs were conducted, separately fitting the same two-factor model to the data collected at T1 and T2. Table 9.9 (p. 189) reports the goodness-of-fit indices for these analyses. As the hypothesised model did not achieve a good fit at either time point, the factor loadings, residual variances, and modification indices were consulted to gain an understanding of where the tested model diverged from the patterns in the collected data. The following stepwise changes were implemented, in the end achieving a good model fit:

1. The item Informal 1 (*'I think that the internet can help you improve your English'*) was removed because its factor loading was substantively lower than that the other informal items. This low factor loading might be because this item specifically asks about beliefs regarding *'online'* language use (whereas the other questions refer more generally to using English *'in [one's] free time'* or *'outside of school'*), thus this item may measure attitudes towards the internet itself in addition to informal learning.
2. The item Formal 4 (*'When learning a language, it's important to practice vocabulary explicitly'*) was similarly excluded due to its lower factor loading, as well as a strong cross-loading onto attitudes towards *informal* learning. This was surprising, given that a very similar item (Formal 2: *'When learning a*

language, it's important to know the grammar rules') did not show such a cross-loading. A possible explanation of this difference can be found by looking back at the focus group data on which the researcher drew to develop these items (Section 5.3): The participants in Phase I seemed to strongly associate focusing on grammar with classroom learning, as illustrated in frequent comments showcasing their belief that formal instruction was necessary to acquire basic grammatical knowledge, or else that they disliked the grammar-focus of their lessons. Vocabulary, on the other hand, was more frequently discussed in the context of *informal* rather than formal learning, with participants claiming that they had learned a substantial part of their vocabulary through ISLPs. This might serve to explain the low loading of the vocabulary-focused item (Formal 4) onto the attitudes towards formal learning factor, and its high cross-loading onto attitudes towards informal learning (which was negative, suggesting that agreement with the importance of explicit vocabulary instruction indicated a lower belief in the effectiveness of informal learning).

The above discussion illustrates that it is difficult to completely disentangle attitudes towards informal and formal learning from each other, since a strong belief in the effectiveness of one seems to relate to believing the other mode of learning to be less effective. This was also implicit in the wording of some items (e.g. Informal 2: '*The English I encounter in my free time is more useful to me than what I learn in school*') and furthermore demonstrated by the strong negative correlation between the two latent factors ($r_{T1} = -.53$ and $r_{T2} = -.47$, see Figures 9.8 and 9.9).

Altogether, excluding the two items discussed above significantly improved the fit of the model ($\Delta\chi^2_{T1}[15] = 78.88, p < .001$; $\Delta\chi^2_{T2}[15] = 120.15, p < .001$), with

all indices (except for the Chi-square statistic) reaching the range of ‘acceptable’ to ‘good’. While the factor loadings of some items were still substantially lower than others, especially at T1, they were clearly above .40, which researchers tend to agree indicates a ‘clear factor loading’ (e.g. Hinkin, 1998; Worthington & Whittaker, 2006).

Figure 9.8: Final two-factor model of attitudes towards language learning at T1

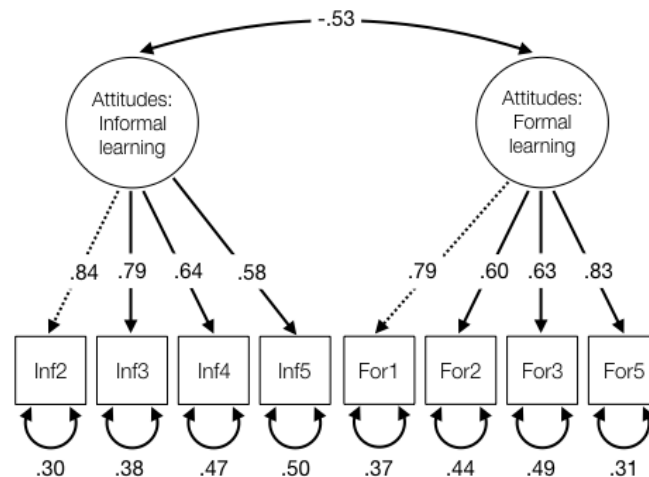
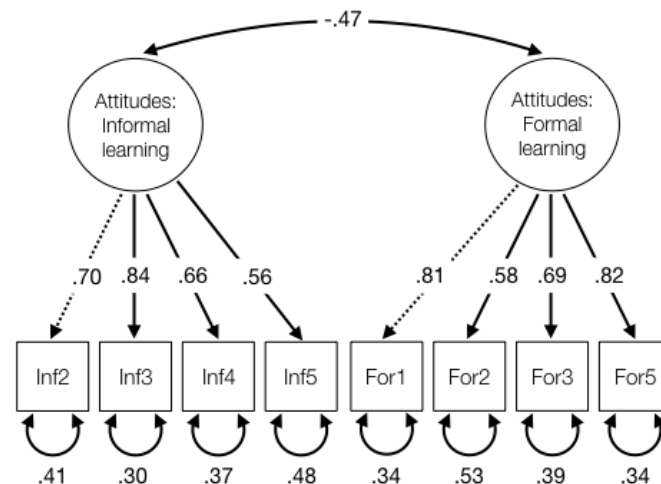


Figure 9.9: Final two-factor model of attitudes towards language learning at T2



Finally, longitudinal MIMs were also conducted for these models (see Table 9.10, p. 189), including correlations between the residuals of the corresponding indicators across the time points (e.g. Formal 1.T1 with Formal 1.T2). According to the Chi-square statistic, the model fit was significantly worse for scalar invariance model than

Table 9.5: Goodness-of-fit indices, CFA of current and future L2 self at T1 and T2

Model	Time 1 (n = 320)							Time 2 (n = 281)						
	χ^2	df	p	CFI	TLI	RMSEA	SRMR	χ^2	df	p	CFI	TLI	RMSEA	SRMR
Baseline (as planned)	166.52	48	< .001	.95	.93	.10	.04	151.96	48	< .001	.96	.94	.10	.06
Excl. Future 2 and Current 2	114.95	30	< .001	.96	.93	.09	.03	78.36	30	< .001	.98	.96	.09	.04

Table 9.6: Goodness-of-fit indices, MIMs of current and future L2 self at T1 and T2

Model	χ^2	df	p	CFI	TLI	RMSEA	SRMR
Configural Invariance	287.25	144	< .001	.97	.96	.06	.04
Metric Invariance	299.17	152	< .001	.97	.96	.06	.05
Scalar Invariance	322.16	162	< .001	.97	.96	.06	.05
Strict Invariance	337.59	172	< .001	.96	.96	.06	.05

Table 9.7: Goodness-of-fit indices, CFA of motivational self-distance at T1 and T2

Model	Time 1 (n = 320)							Time 2 (n = 281)						
	χ^2	df	p	CFI	TLI	RMSEA	SRMR	χ^2	df	p	CFI	TLI	RMSEA	SRMR
Baseline (as planned)	54.52	9	< .001	.94	.92	—	.06	27.37	9	< .001	.98	.96	—	.03
Correlated residuals Distance5, Distance6	37.66	8	< .001	.97	.95	—	.03	22.26	8	< .001	.98	.96	—	.03

Note: The RMSEA was not consulted in these analyses due to its unreliability in models with fewer degrees of freedom.

Table 9.8: Goodness-of-fit indices, MIMs of motivational self-distance at T1 and T2

<i>Model</i>	χ^2	<i>df</i>	<i>p</i>	<i>CFI</i>	<i>TLI</i>	<i>RMSEA</i>	<i>SRMR</i>
<i>Configural Invariance</i>	98.62	45	< .001	.98	.97	.07	.03
<i>Metric Invariance</i>	110.34	50	< .001	.97	.97	.07	.05
<i>Scalar Invariance</i>	116.87	56	< .001	.97	.97	.06	.05
<i>Strict Invariance</i>	119.22	62	< .001	.97	.97	.06	.05

Table 9.9: Goodness-of-fit indices, CFA of attitudes towards informal and formal language learning at T1 and T2

<i>Model</i>	<i>Time 1 (n = 319)</i>							<i>Time 2 (n = 281)</i>						
	χ^2	<i>df</i>	<i>p</i>	<i>CFI</i>	<i>TLI</i>	<i>RMSEA</i>	<i>SRMR</i>	χ^2	<i>df</i>	<i>p</i>	<i>CFI</i>	<i>TLI</i>	<i>RMSEA</i>	<i>SRMR</i>
<i>Baseline (as planned)</i>	126.83	34	< .001	.88	.84	.10	.08	177.09	34	< .001	.82	.76	.12	.10
<i>Exl. Formal 4</i>	93.60	26	< .001	.90	.86	.10	.08	104.25	26	< .001	.88	.84	.11	.09
<i>Excl. Informal 1</i>	47.95	19	< .001	.95	.93	.08	.06	56.94	19	< .001	.94	.91	.09	.07

Table 9.10: Goodness-of-fit indices, MIMs of attitudes towards informal and formal language learning at T1 and T2

<i>Model</i>	χ^2	<i>df</i>	<i>p</i>	<i>CFI</i>	<i>TLI</i>	<i>RMSEA</i>	<i>SRMR</i>
<i>Configural Invariance</i>	152.05	90	< .001	.97	.95	.05	.06
<i>Metric Invariance</i>	164.01	96	< .001	.96	.95	.05	.06
<i>Scalar Invariance</i>	212.11	104	< .001	.94	.93	.06	.07
<i>Strict Invariance</i>	216.58	112	< .001	.94	.94	.06	.07

the metric invariance model ($\Delta\chi^2_{scal,metr}[8] = 46.10, p < .001$). Yet, the most restrictive invariance model still achieved an ‘acceptable’ fit across all of the other indicators, supporting the validity of modelling changes in attitudes across the two time points as planned.

9.1.3.5 Engagement

Because of the unique method of collecting the engagement data, a different analytical was necessary for the analysis of this data: While the participants’ proficiency, their motivation, and attitudes towards language learning were measured at two time points, data about the participants’ engagement in ISLL stemmed from diary entries recorded over the course of several days, one after each instance of informal L2 contact. Essentially, the engagement data were *hierarchical*, meaning that a variable number of individual data points (diary entries) are nested within a higher-order structure (participants). In statistical terms, this means that the collection of an individual participant’s engagement scores across all submitted diary entries is expected to reflect some activity-specific variance, as well as some person-level variance (i.e. an individual’s *general* levels of engagement in ISLPs).

In the context of this study — which aimed to explore variations in engagement *between* (as opposed to *within*) participants and how they relate to proficiency, motivation, and attitudes — it was important to derive accurate values to represent each person’s *general* levels of engagement. Calculating an average score for each engagement item across all diary entries (as had been done in Phase II, see Section 7.3), while the easiest way to aggregate the available data, is problematic, because it may be biased by the within-participant variance. A more accurate approach is to use *Multi-level Confirmatory Factor Analysis* (MCFA) to separately

analyse engagement at the activity- and participant-level. This not only provides less biased measures of participants' general engagement (because the per-participant aggregate values have been adjusted for the variation in engagement between different activities), but also allows for the investigation of how the different dimensions of engagement investigated in this study interrelated both at the individual activity- and the participant-level.

During the EFA in Phase II (Section 7.3), affective, cognitive, and linguistic engagement emerged as three separate factors, as hypothesised. However, given that linguistic engagement is proposed in this study as a new and theoretically separate dimension from cognitive engagement, the researcher decided to use the larger data set gathered in Phase III to conduct further analyses on the distinctiveness of these two factors. Two initial MCFAs were therefore conducted with the nine engagement items, to compare the fit of the two- and three-factor models (see Appendix E.5 and Table 9.11 below for goodness-of-fit indices).

Table 9.11: Goodness-of-fit indices, MCFAs of two- and three-factor engagement models

<i>Model</i>	χ^2	<i>df</i>	<i>p</i>	<i>CFI</i>	<i>TLI</i>	<i>RMSEA</i>	<i>SRMR_{w/i}</i>	<i>SRMR_{b/w}</i>
<i>Two-factor model</i>	630.66	52	< .001	.80	.72	.10	.10	.25
<i>Two-factor model, excluding Language 3</i>	436.37	38	< .001	.87	.82	.10	.10	.20
<i>Three-factor model</i>	502.24	48	< .001	.85	.78	.09	.10	.25
<i>Three-factor model, excluding Language 3</i>	297.26	34	< .001	.92	.87	.08	.07	.19
<i>Three-factor model, letting Cognition 2 load only onto linguistic engagement</i>	202.69	34	< .001	.97	.95	.05	.04	.05
<i>Three-factor model, cross-loading Cognition 2</i>	220.02	32	< .001	.97	.95	.05	.04	.06

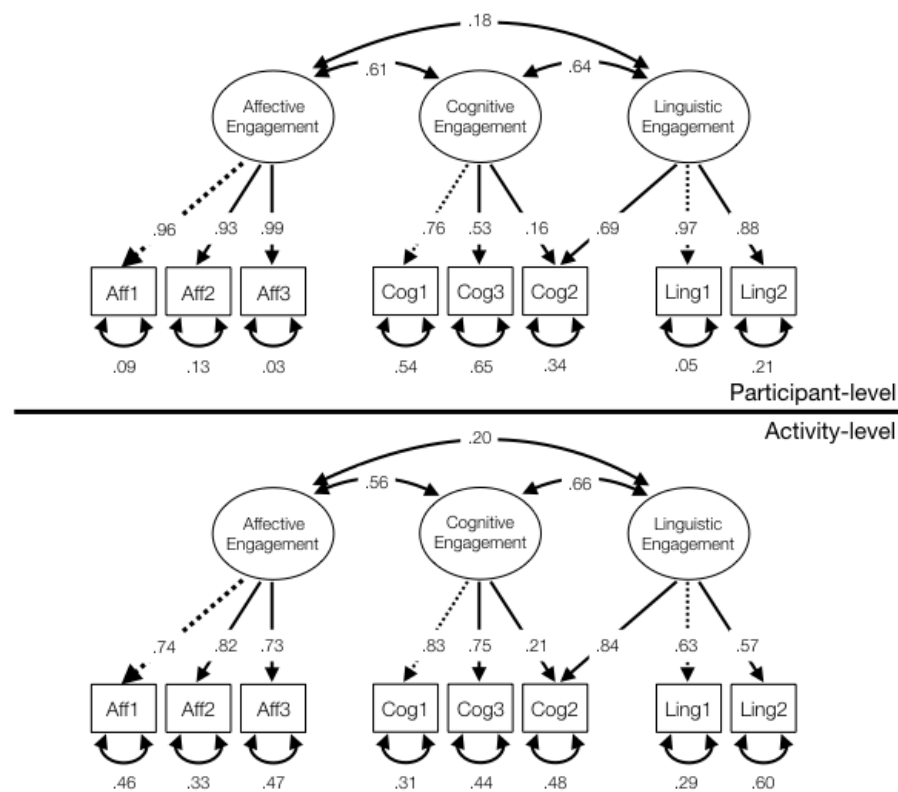
Number of observations used in these analyses: 1505 entries by 254 participants.

Although the fit indices for the three-factor model were substantively better than for the two-factor model ($\Delta\chi^2[4] = 128.42, p < .001$), at first neither model achieved an ‘acceptable’ fit according to any of the fit indices. To identify the source of this misfit, the modification indices (MI) and item factor loadings were consulted. In both the two- and three-factor models, the Language 3 item (*‘During this activity, I focused on [the general meaning – individual words/phrases]’*; which had been added only after the piloting phase) stood out by its extremely low loading onto the linguistic engagement factor at the activity-level (.03 and .15 respectively) and negative loading at the participant-level. The modification indices also suggested a high cross-loading of this item onto the factor representing affective engagement. Clearly, this item was responsible for a large part of the model misfit and did not measure linguistic (or in the two-factor model, *cognitive-linguistic*) engagement, as intended. Both the two- and three-factor CFA were therefore re-run while excluding this item. Again, the three-factor model achieved a significantly better fit than the two-factor model ($\Delta\chi^2[4] = 139.11, p < .001$). This and the overall improved factor loadings, which were higher in the three-factor than the two-factor model, suggested a stronger structural validity of the three-factor solution (see Figure 9.10 below and Figure E.10 in Appendix E.5).

Yet, some fit indices of the three-factor model were still not within the range indicative of an ‘acceptable’ model fit (especially the low TLI and high participant-level SRMR). The MI suggested that this was because the item Cognition 2 (*‘I was [not at all – completely] focused on trying to understand every single word’*) exhibited a substantive cross-loading onto linguistic engagement. This was not altogether surprising, given the item wording and the fact that it had initially been designed to measure linguistic engagement (it emerged in the EFA in Phase II as loading more strongly onto cognitive engagement, see Section 7.3). Allowing this item to cross-load

on both factors in the current MCFA significantly improved the fit as compared to the model without the cross-loading ($\Delta\chi^2[2] = 77.24, p < .001$). The researcher also tried to let this item load only onto linguistic engagement instead. This did not fundamentally improve the overall model fit but did markedly decrease the factor loadings of the other cognition indicators and increased the correlation between linguistic and cognitive engagement. It was therefore decided that cross-loading the item onto both factors was the better solution, increasing the validity of the framework enough to justify the added model complexity (cf. Byrne, 2010).

Figure 9.10: Final three-factor model of engagement at the activity- and participant-level



Effects of missing engagement data

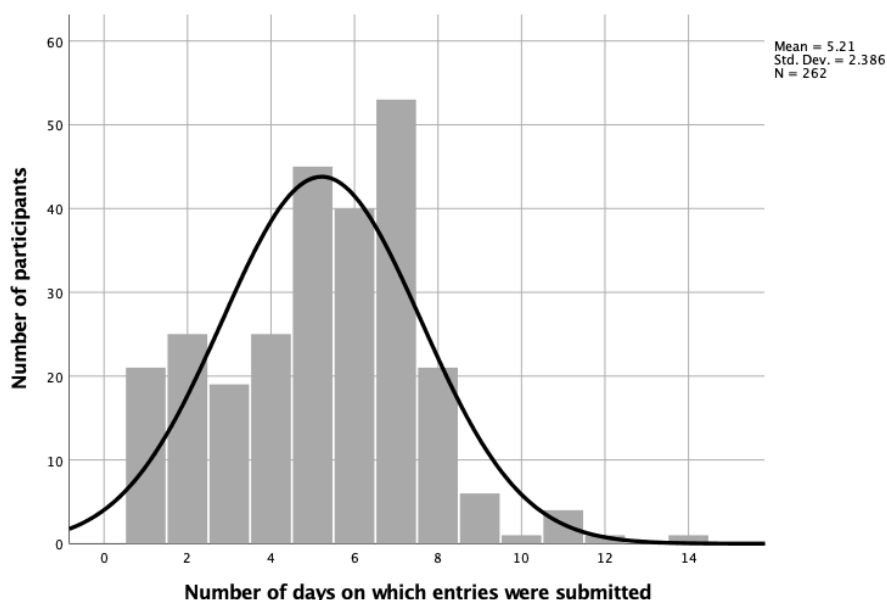
At this point, another aspect to the validity of the engagement scores must be considered: Even though the participants had been instructed to use the informal language learning diary to record their ISLPs for seven days, there was substantial individual variation in the number of days across which individual pupils actually submitted entries (Figure 9.11), as indicated by the time stamps in the survey system. The majority of the students (60.5%) recorded activities across five to eight days, while some (8%) submitted entries for only one day and others (5%) kept the diary for as long as nine to fourteen days. Therefore, it must be investigated to what extent the fit indices, factor loadings, and error variances discussed above (which supported the validity of the engagement instrument) might be biased by data from participants who submitted only a few diary entries, across a small number of days.

To answer this question, the researcher ran the MCFA twice more, excluding participants who made entries over fewer than four or five days (respectively: $n = 64$ and $n = 91$). While the model did have better fit for the reduced data sets, at least in terms of the Chi-square statistic (Table E.11, Appendix E.5), the factor loadings, error variances, and inter-factor correlations did not change substantively (Figures E.11 and E.12, Appendix E.5). Additionally, an analysis was conducted which included the number of days across which a participant submitted diary entries as a predictor of affective, cognitive, and linguistic engagement at the participant level: This revealed that the number of days did not, in fact, significantly predict changes in any of these three variables (*affective engagement*: $B = .00, SE = .04, p = .93$; *cognitive engagement*: $B = .06, SE = .04, p = .11$; *linguistic engagement*: $B = .02, SE = .04, p = .62$). These analyses suggest that the MCFA model of engagement

presented above was not substantially biased by including the data from students who submitted only a few diary entries.

Altogether, the analyses discussed in this section support the validity of the items included in the informal language learning diary (with the exception of Language 3) as measures of the three unique aspects of engagement, even when including cases where only small amounts of data were available per participant.

Figure 9.11: Number of days for which participants submitted diary entries



9.1.3.6 Factor Scores

Besides providing a way to evaluate the validity of quantitative instruments, CFAs, can also be used to derive *factor scores* (FS): aggregate values, weighted according to the individual items' factor loadings, which can be used to represent the latent constructs in subsequent analyses (DiStefano, Zhu, & Mîndrilă, 2009). FS for all proficiency, motivation, attitude, and engagement variables discussed above were

produced in lavaan, for use in some of the later LPA and SEM analyses (Sections 9.2–9.5). It should be noted that the software automatically standardises factor scores to a mean of 0 and a variance of 1, which can make them difficult to interpret in relation to the original measurement instruments.

FS for proficiency, motivation, and attitudes at T1 and T2 were produced using the longitudinal metric invariance models (i.e. imposing the same factor structure and factor loadings at both time points). The affective, cognitive, and linguistic engagement factor scores (both at the activity- and participant-level) were derived from the final three-factor multi-level model.

9.1.4 Reliability: Cronbach's Alpha

While the CFAs and MIMs discussed above provided support for the validity of the instruments developed as part of this study, reliability analyses were carried out again at this point to confirm the reliability of the final (in some cases reduced) groups of items. The results are presented in Table 9.12.

The majority of the scales had good internal consistency at both time points (all Cronbach's $\alpha > .80$). The attitudes towards informal and formal learning, instrumental benefits, knowledge orientation, international orientation, and cognitive engagement scales had slightly lower alpha scores, although they were still well within the range of acceptability for Social Science research (all Cronbach's $\alpha > .70$; see Kline, 2000). For all scales, the item-total correlations indicated that all items correlated sufficiently with the scales as a whole ($r > .40$; see Blunch, 2013) and the values for Cronbach's alpha if any *one* of the items were to be deleted

confirmed that the internal consistency of none of them would be improved by excluding any one item.

Table 9.12: Reliability analyses, Phase III

<i>Scale</i>	<i>Cronbach's alpha</i>
C-TESTS	
Version 1, T1	.90
Version 1, T2	.81
Version 2, T1	.92
Version 2, T2	.93
LANGUAGE LEARNING ORIENTATIONS	
Intrinsic motivation, T1	.89
Intrinsic motivation, T2	.88
Instrumental motivation (obligation), T1	.83
Instrumental motivation (obligation), T2	.84
Instrumental motivation (benefits), T1	.79
Instrumental motivation (benefits), T2	.80
Knowledge orientation, T1	.74
Knowledge orientation, T2	.73
International orientation, T1	.76
International orientation, T2	.77
SECOND LANGUAGE MOTIVATIONAL SELVES	
Ideal future self, T1	.90
Ideal future self, T2	.90
Current self, T1	.94
Current self, T2	.94
ATTITUDES TOWARDS LANGUAGE LEARNING	
Attitudes towards informal learning, T1	.76
Attitudes towards informal learning, T2	.76
Attitudes towards formal learning, T1	.71
Attitudes towards formal learning, T2	.74
ENGAGEMENT	
Affective engagement, T1	.90
Cognitive engagement, T2	.77
Linguistic engagement, T1	.83

9.2 Research Question 1: How do German Learners of English as a Second Language Attending *Gymnasien* Engage in English Informal Second Language Practices?

In addressing this first research question, a Multi-Level Confirmatory Factor Analysis as well as descriptive and basic inferential statistical methods were used to explore the data collected through the informal language learning diary.

The first analyses were to determine how much time the pupils in this study spent on informal second language practices (ISLPs), how varied their repertoires of informal practices were (both aspects of behavioural engagement; Section 9.2.1), which activities were most popular (Section 9.2.2), and whether there were differences in the levels of affective, cognitive, and linguistic engagement associated with different informal practices (Section 9.2.3). Next, the an MCFA was used to explore the relationships between the different dimensions of engagement that were measured in this study (behavioural, affective, cognitive, and linguistic engagement; Section 9.2.4). Finally, it was investigated whether groups of learners could be identified who showed distinct overall patterns of engagement in informal L2 practices (Section 9.2.5).

9.2.1 Quantity and Variety of Informal Second Language Practices

This section introduces the per-participant measures of behavioural engagement (the quantity and variety of participants' ISLPs) use in the later SEM analyses. As mentioned previously (Section 9.1.3.5), participant-level affective, cognitive, and linguistic engagement factor scores were derived from the Multi-Level CFA model.

To measure the quantity of participants' ISLPs, it was not enough to take a sum of the durations of all activities reported by each of the participants. This was because, as

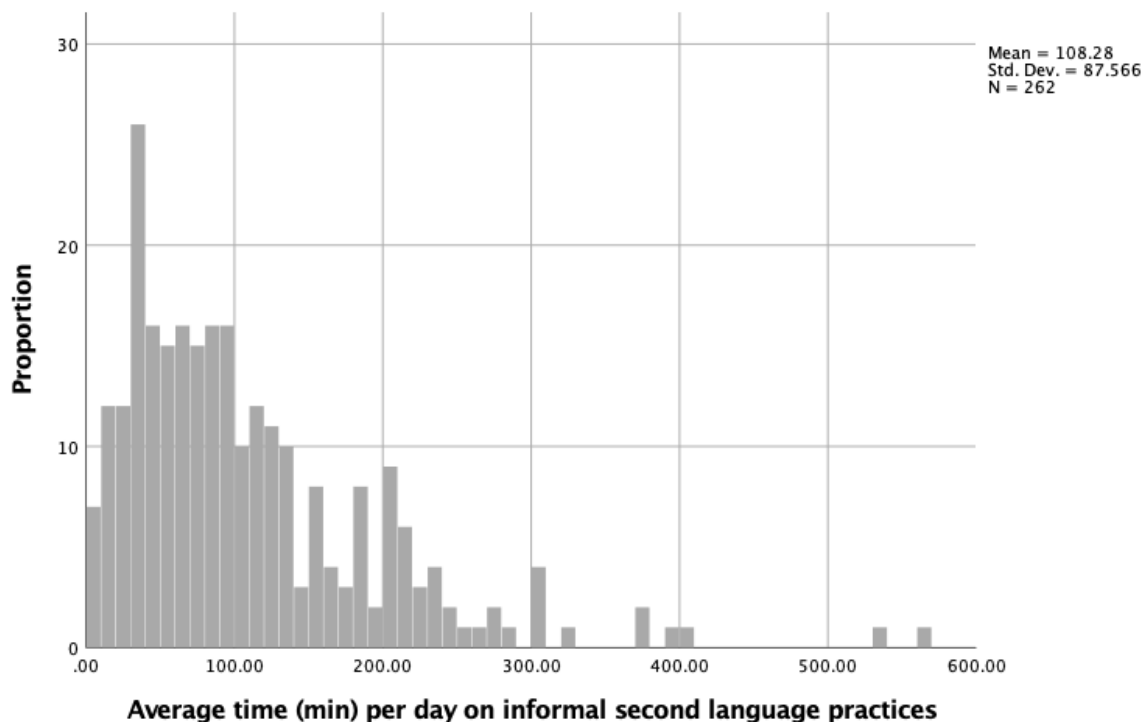
mentioned previously, there was substantial individual variation in the number of days across which each participant submitted diary entries. To account for this, the durations of each participant's activities were summed per day and then averaged over the number of days for which data were available (to obtain per-day averages for each participant).

As can be seen in Figure 9.12, the distribution of these scores showed a slight positive skewness and very high kurtosis ($S = 1.80$; $K = 4.86$), meaning that most of the participants' scores clustered at the lower end of the scale. The majority of the participants (70.3%) engaged in informal practices for less than 2 hours per day: 13.7% of the students spent between 0 and 30 min daily on ISLPs, 23.3% between 31 and 60 min, and 33.3% between 61 and 120 min (1–2 hours). A further 13.3% of the scores fell between 120 and 180 minutes (2–3 hours), 5.1% between 181–240 min (3–4 hours), and finally 6.5% above 240 (more than 4 hours). Across the whole sample, the mean time spent on ISLPs per day was 1.8 hours (108.3min), with considerable individual variation ($SD = 87.57\text{min}$).

There were a few outlying values at the higher end this range, which also help to explain the non-normality of the overall data distribution. The highest value (560min) came from a participant who 'bulk reported' all of his daily informal activities in one diary entry per day (across a total of three days), instead of making separate entries for each activity as instructed. Thus, it seems likely that this was not a very accurate estimate of how much time this participant actually spent on informal activities per day. This data point was thus excluded from the further analysis. The second highest value (539min) stemmed from a participant who submitted 46 diary entries over the course of 11 days, and thus seemed reliable enough to be included in the data set.

It must be recognised that the validity of these values will depend heavily on how representative the recorded days' activities were of the students' general informal second language practices, especially if an individual pupil only submitted diary entries across a small number of days. Reassuringly, the ten participants who had the lowest and highest values for time spent on ISLPs per day (after excluding the outlier discussed above) recorded data across at least three days. This shows that the overall distribution was unlikely to be influenced heavily by outlying values belonging to participants who recorded only a single day of unusually short or long engagement in ISLPs.

Figure 9.12: Behavioural engagement (quantity) distribution of scores



The variety of the participants' informal practices, in terms of the number of different activities with which each of them engaged over a given period of time, is more difficult to operationalise than the quantity, given that scores will depend heavily on

the way in which activities are categorised as the *same* or *different*. The general categories that will be used in some of the later analyses (Sections 9.2.2 and 9.2.3) were chosen because they make the results comparable to earlier research on ISLL. However, they are still likely to be too broad to accurately capture the variation in individual ISLP repertoires.

There are some previous studies which have also sought to measure variety in informal second language practices: Lai, Zhu and Gong (2015) used open-ended questionnaires and focus group interviews to gather information about the quantity and 'diversity' of their participants' ISLPs. These researchers relied on the participants' own perceptions in determining what to count as distinct informal learning activities:

For instance, reading newspapers and reading novels were counted as two activities because, although both activities involve reading, the fact that the participants had noted them as two entries in their response suggested that these activities were perceived as different activities and might have served different functions for the participants. (Lai et al., 2015, p. 286)

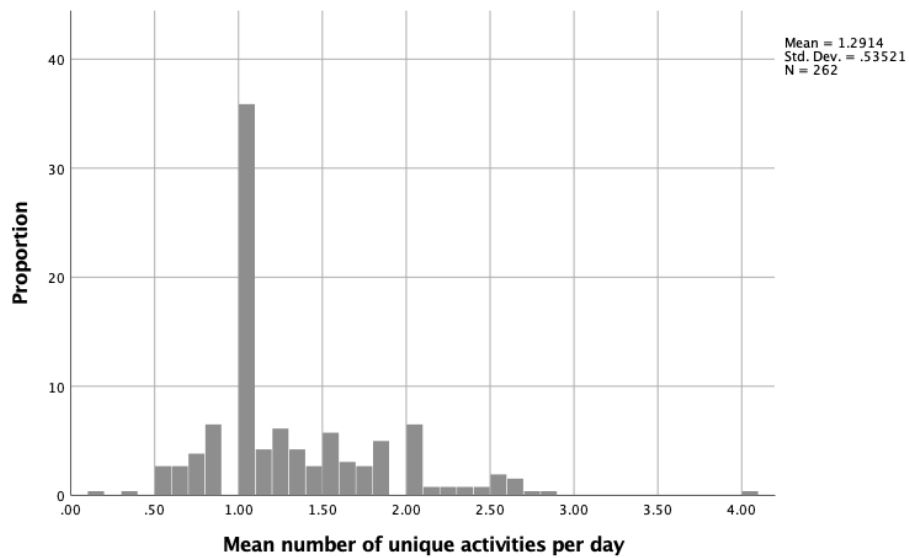
The same procedure was adopted in later studies which also measured variety in ISLPs (J. S. Lee, 2017; J. S. Lee & Dressman, 2017). Unfortunately, it was not practical to use this method in the current study, where the behavioural engagement data were collected through a closed- rather than open-ended questionnaire, to allow for easier processing of data from a larger sample. That is, the participants were asked to choose from a list of activities which one they had engaged in and were asked specific follow-up questions depending on their selection (Appendix B.3.4). From this data, it was

possible to derive a comprehensive list of 31 specific activity categories (Appendix F.1), which were then used to calculate the average number of unique practices in which each participant engaged which served as a measure of the variety in their ISLPs.

It is worth noting that the 31 specific activity types, while based on the data from the closed-ended questionnaires, were not imposed by the researcher, but formulated (in the spirit of Lai et al.'s [2015] approach) based on the data from the focus group interviews with participants from the same population (Phase I, Section 5.1). Additionally, the students who completed the diary also had the option of recording an '*other*' activity, with a brief description of what they did. 42 such entries were submitted: 16 were instances where participants trying to 'bulk report' all of their informal practices from one day in a single diary entry, rather than making separate entries for each activity; six were cases of participants trying recording a *formal* practice (i.e. studying for an English exam); and of the rest, all but two entries (reports of using English for computer programming) could accurately be re-coded into one of the pre-determined activity categories. Thus, it can be argued that the options provided in the survey were sufficiently diverse to accurately capture the variety of the participants' ISLPs.

As with the quantity, there was a need to account for the variability in the number of days across which individuals participated in the informal language learning diary when calculating the measure of the variety of ISL activities. The number of each participant's different activities were thus counted per day and then averaged over the number of days for which data was available (to obtain per-person, per-day averages; Figure 9.13).

Figure 9.13: Behavioural engagement (variety) distribution of scores



Across the whole sample, the mean number of unique informal activities per day was 1.29 ($SD = .54$). The lowest value was .14 and the highest 4.20. As can be seen in Figure 9.14, a large proportion of the scores (35.5%) clustered around 1, which is reflected in the strong positive kurtosis ($K = 2.31$; $S = 1.26$). It must be kept in mind that these low values do not mean that the participants engaged in the same ISLPs every day, but rather that there was little variety *within one day* for a large proportion of the sampled students.

As with the quantity of ISLPs, the validity of the variety measure depends heavily on how representative the recorded days' activities were of an individual's general informal practices, especially if they submitted only a small number of diary entries across a few days. Reassuringly, the ten values at the lower and upper ends of the distribution stemmed from participants who recorded diary entries across at least four days. Again this means that the overall distribution was unlikely to be influenced extremely by outlying values from participants who only recorded a single day of unusually diverse or limited engagement in ISLPs.

9.2.2 Popularity of Different Informal Second Language Activities

In each diary entry, the participants reported the general type of English-language activity in which they had engaged: watching films, serials and TV shows, engaging with social media, writing or speaking in English, listening to music, playing video games, reading, or listening to spoken English. Looking at these particular categories means that their popularity among this sample could be compared to earlier studies of different populations (see Section 2.1.1). However, these categories are relatively broad and do not always give an accurate idea of the ways in which the participants in this study actually *used* the English language in these practices. For example, *using social media* can encompass a wide variety of behaviours, including watching videos, reading texts of varying lengths, and even producing written or spoken language. For this reason, the participants were also asked a number of follow-up questions in the informal language learning diary, determined by which type of activity they were recording. This not only helped the researcher to better understand the nature of the participants' ISLPs, but it also facilitated coding the entries into more specific sub-categories for use in some of the later analyses (see Section 9.2.1).

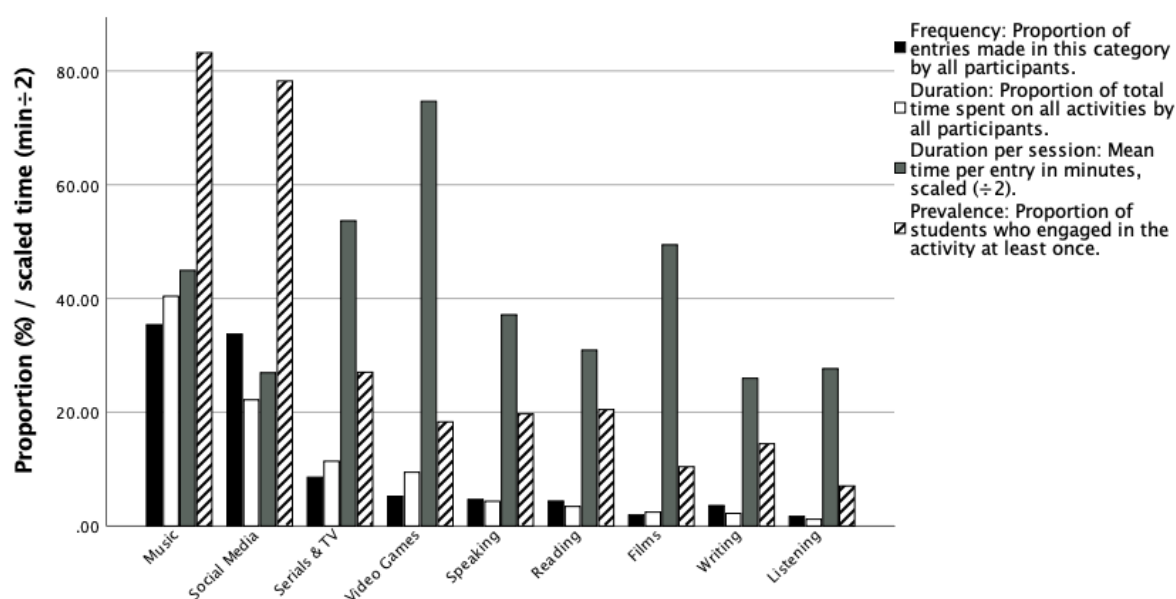
In this section, some of the main findings regarding the relative popularity of the broader activity categories are laid out. A more detailed discussion of each activity type and its sub-categories can be found in Appendix F.3.

The relative popularity of different activities among this sample of participants was explored using four criteria (see Figure 9.14 below and Appendix F.2):

- (1) **Frequency:** how often participants engaged with each type of activity, as measured by the number/proportion of diary entries in this category;
- (2) **Duration:** how much time participants spent on this type of activity, measured by

the total duration of all entries in this category; (3) **Duration per session**: the mean time recorded per diary entry in this category; and (4) **Prevalence**: how many participants engaged with each type of activity, that is, submitted at least one diary entry in this category.

Figure 9.14: Relative popularity of informal activities according to frequency, duration, duration per session, and prevalence



All in all, the conclusions that can be drawn about the relative popularity of different activities according to the frequency, total duration, and prevalence criteria correspond well with one another: That is, the activities which accounted for the largest proportion of the recorded total time spent on ISLPs (listening to music, browsing social media, and watching serials) were also those with which the participants engaged most frequently and in which the largest proportion of participants engaged at least once. Nevertheless, a few interesting findings emerged from comparing rankings across the different criteria: For example, while listening to music and browsing social media were comparable in terms of frequency and prevalence, the participants as a whole spent almost twice as much time on music.

The fact that music outranked all other activities by such a large margin, in terms of total time spent on it, is most likely due to the combination of a high frequency (large number of entries) and the relatively long duration per session (students listened to music for long stretches of time).

In fact, the duration per session provided a noticeably different picture of informal second language practices within this sample than the other criteria: Whereas music and social media accounted for the majority of the total time spent on ISLPs, participants spent by far the longest *continuous* periods of time playing video games ($M = 149.3$ min, $SD = 102.5$) and watching serials ($M = 107.3$ min, $SD = 76.1$). Of course, the vast individual variation (expressed by the extreme sample standard deviations) must be taken into consideration here, as well as the possibility of concurrent activities: For example, participants could have listened to music while playing a video game or used social networking sites on their phone while also watching serials on their computer or television.

Similar to the duration per session, the prevalence criterion (the proportion of students who reported engaging with an activity at least once) also added an interesting dimension to the relative popularity of ISL activities, due to the very large discrepancies between different activities: While the vast majority of students reported, at some point during the study, to have listened to music (83.2%) or used social media in English (78.2%), other activities were not nearly as widespread: Only about 27% of the participants watched any serials or listened to spoken English and no more than 20% reported ever speaking English, playing video games, writing, or watching films in English.

9.2.2.1 *Informal Second Language Practices and Gender*

Some prior ISLL studies have reported significant differences between the types of activities that are most popular among boys and girls (e.g. Kuppens, 2007, 2010; Olsson, 2011; Sundqvist & Sylvén, 2014). To determine whether such gender differences existed in the current sample, the researcher planned to compare the average time that boys and girls spent per day on thirteen different types of informal activities (listed in Table 9.13, p. 213). However, it was not possible to conduct a simple comparison of the group means because the distribution of these scores was heavily skewed: For each activity, there were large numbers of participants who did not engage in it at all, that is, spent zero minutes per day on this activity (see ‘prevalence of different activities’, Table F.2, Appendix F.2).

This type of data distribution is called ‘zero-inflated’ and a two-step approach was adopted to deal it: First, a series of logistic regression analyses was used to test whether gender could predict a set of thirteen binary variables, each coding for whether participants had engaged in a particular type of activity (=1) or not (=0). Subsequently, all participants who had not engaged in the activity were excluded (i.e. all zero-scores recoded into missing values), and a series of simple linear regressions was used to test whether gender predicted the average amount of time that individual participants spent the different activities. Instead of conducting separate logistic and simple regression analyses for each activity type, it would have been preferable to use two path models to simultaneously estimate the effect of gender on all activity types (one model for the binary outcome variables, one for the continuous one). This is because in the simultaneous analysis, the regression coefficients would have been adjusted for any covariances between the outcomes. However, due to the large number of missing data in the continuous variables

representing time spent on each activity, not enough data was available for this model to converge. Hence, separate analyses were run for each activity, the results of which are reported in full in Table F.6, Appendix F.4 and summarised below.

In the logistic regressions, gender significantly predicted of the likelihood of engaging in all types of activities except for purely video-based and text-based social media (all $p < .05$). Boys were more than 2.5 times as likely as girls to engage in playing video games and 15 times as likely to play video games while also communicating with other players. Gender explained 5% and 17% of the variance in these two variables respectively (Nagelkerke r^2 , small and medium-sized effects).

Girls emerged as significantly more likely to engage in all other activities than boys. This gender effect was strongest for watching films and listening to music, which girls were respectively 6.6 and 3.3 times as likely to do. Gender explained, respectively, 14 and 11% of the variance in these two variables respectively (Nagelkerke r^2). The effects on the other activities were much smaller, with girls being only between 1.6 times (listening to music in the background) and 2.5 times (watching serials, speaking) as likely to engage in them (r_N^2 between .03 and .05).

While boys and girls were thus likely to engage in different activities, gender significantly predicted the average *time spent per day* on only one activity: using video-based social media ($B = -21.77$; $SE = 5.35$; $\beta = -.35$; $p < .001$). The male participants in this study spent significantly more time on this activity ($M = 36.88$; $SD = 46.85$) than the girls ($M = 15.19$; $SD = 17.00$), with gender explaining 11% of the variance in this variable.

Overall, then, video-based social media were more popular among boys in this sample, whereas more girls used mixed social media, which involved both reading

and watching videos. An explanation for these contrasting findings can be found in the specific social media platforms that were more popular among both genders: It was determined that the majority of video-based social media engagement happened on YouTube (79.5% of all entries), followed by Instagram (18.6%) and other platforms (1.9%). Meanwhile, mixed social media engagement was split almost equally between Instagram (47.0% of entries) and YouTube (43.8%). This suggests that the aforementioned gender differences might be caused by differences between boys' and girls' use of two primary platforms, YouTube and Instagram. Indeed, further regressions analyses revealed that girls were almost three times as likely as boys to use Instagram ($B = -1.07, SE = .25, p < .001$; no significant differences for YouTube), and thus to engage in mixed social media consumption. Meanwhile, boys spent significantly more time on YouTube, and thus on video-based social media ($B = -16.40; SE = 5.07; \beta = -.27; p < .01$; no significant differences for Instagram). Both of these represented small gender effects on the choice of social media platform ($r_N^2 = .08$ and $.07$) and serve to explain the gender differences regarding video-based and mixed social media use presented above.

While some interesting gender differences were revealed with regard to the types of informal activities in which boys and girls engaged, conducting separate analyses by gender for each of the subsequent research questions (Sections 9.3.–9.5) would have gone beyond the scope of this thesis, in terms of the size of the data set required for such analyses and the available project time. Nevertheless, the researcher wanted to know whether there were any significant gender differences in any of the aspects of engagement in ISLPs which would suggest that the variable should be included in the subsequent models as a covariate. Thus, a series of Mann-Whitney U tests was

conducted which compared the distribution of the five factor scores representing affective, behavioural, cognitive, and linguistic engagement, as well as the scores for proficiency at both points of data collection, across the male and female participants. No significant differences were found in any of these variables (all $p > .01$, α adjusted for multiple comparisons).

9.2.3 Levels of Engagement in Different Informal Second Language Activities

In the hope of better understanding the different forms that engagement in various activities took for these participants, a series of Kruskal-Wallis tests were conducted in which the average levels of engagement for each activity type were compared in terms of four statistics (Table 9.13, p. 213): The average duration per session (in minutes), representing behavioural engagement, and the three MCFA-derived participant-level factor scores for affective, cognitive, and linguistic engagement. The Kruskal-Wallis test was chosen, non-parametric alternative to the factorial ANOVA, due to the very large standard deviations and skewed distributions of the dependent variables within the activity types, as well as the vastly different numbers of data points per category (ranging from 35 entries for watching films to 423 entries for focused music listening). The main findings of these analyses are explained in this section; a more detailed discussion of what these data showed with regard to individual types of activities can be found in Appendix F.3.

There were significant overall differences between the different activities in the associated levels of behavioural ($\chi^2[12] = 302.22, p < .001$), affective ($\chi^2[12] = 206.93, p < .001$), cognitive ($\chi^2[12] = 465.29, p < .001$), and linguistic engagement ($\chi^2[12] = 290.03, p < .001$). Thus, stepwise step-down post-hoc analyses were employed to investigate which activities differed significantly from

each other in which aspect. The results are coded by the superscripts in Table 9.13, with identical letters marking subsets of activities which did *not* significantly differ from each other and different letters marking sets between which significant differences existed.

The behavioural engagement (duration per session) was significantly higher for *playing video games while chatting with other players* than for any other activity, followed by just playing video games, watching films, and watching serials. Significantly shorter continuous periods of time were spent listening to music (in the fore- or background), and yet again on speaking, video-based and mixed social media. The lowest durations were recorded for text-based social media and writing.

In terms of affective engagement, most categories did not significantly differ from each other, as indicated by the large number of activities in Table 9.13 bearing the same superscripts. However, the analysis clearly indicated that playing video games while communicating with other players, once again, significantly out-ranked all other activities, followed by watching serials. This indicates that the participants felt significantly more emotionally invested in these activities. Listening to music in the background and listening to spoken English, on the other hand, elicited significantly lower levels of affective engagement, followed by writing, text-based social media use and focused music consumption.

The post-hoc tests also revealed that participants recorded significantly lower levels of cognitive engagement while playing music in the background than when they were focusing on it (as could be expected by the definition of these categories). Text-based and mixed social media use, video gaming, and listening to spoken English were associated with significantly higher levels of cognitive engagement than either

type of listening to music. The overall highest levels of cognitive engagement, however, were associated with reading, together with the only three activity categories which included productive language use: speaking, writing, and chatting with other players while gaming.

Interestingly, writing and speaking were also the two categories with the highest levels of linguistic engagement (writing significantly above speaking), whereas *playing video games and chatting* was among the activities with the lowest levels of linguistic engagement (only ahead of listening to music, both in the back- and foreground). This illustrates the value of considering the strength of participants' general cognitive focus (cognitive engagement) separately from their focus on language features (linguistic engagement; see also Section 9.1.3.5). Besides these activities, which elicited particularly a strong or weak focus on language, most other activities did not significantly differ from one another in terms of the associated levels of linguistic engagement.

Table 9.13: Sample-wide levels of engagement per activity type

Activity	Behavioural engagement Duration per Session: Time per entry in minutes					Affective engagement Standardised Factor Scores		Cognitive engagement Standardised Factor Scores		Linguistic engagement Standardised Factor Scores	
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Mean rank</i>	<i>Mdn</i>	<i>Mean rank</i>	<i>Mdn</i>	<i>Mean rank</i>	<i>Mdn</i>	<i>Mean</i>	<i>Mdn</i>
1. Films	35	98.91	57.55	1190.97 ^{c d}	210	865.49 ^{c d}	.205	863.14 ^{c d}	-.255	831.80 ^{c d}	-.256
2. Serials & TV shows	155	107.34	76.12	1175.74 ^d	90	983.36 ^d	.271	932.94 ^{c d}	.266	783.04 ^{c d}	.046
3. Video-based social media	262	60.49	63.49	748.17 ^b	40	844.31 ^c	.152	917.97 ^{c d}	.269	836.17 ^{c d}	.091
4. Text-based social media	77	31.36	27.91	486.68 ^a	30	646.33 ^b	-.066	833.23 ^c	.218	830.85 ^{c d}	.058
5. Mixed social media	256	53.26	49.13	729.57 ^b	30	754.28 ^{b c}	.103	841.14 ^c	.195	823.01 ^{c d}	.089
6. Writing	65	50.27	75.93	564.44 ^a	30	607.74 ^b	-.077	1017.09 ^d	.502	1166.57 ^f	.653
7. Speaking	84	74.25	88.65	763.92 ^b	30	725.88 ^{b c}	.123	958.65 ^d	.374	1068.22 ^e	.554
8. Music (background)	218	97.77	92.89	1007.85 ^c	60	405.26 ^a	.380	215.05 ^a	-.846	332.27 ^a	.508
9. Music (focused)	428	85.92	77.04	981.29 ^c	60	721.04 ^b	.028	562.87 ^b	-.176	636.87 ^b	.106
10. Video games	42	123.50	84.07	1210.21 ^d	113	921.97 ^{c d}	.256	769.59 ^c	.060	702.81 ^{b c}	.054
11. Video games & text or voice chat	52	170.00	111.69	1440.49 ^e	150	1101.10 ^e	.366	976.88 ^d	.346	674.41 ^{b c}	.111
12. Reading	79	62.22	86.03	670.12 ^{a b}	30	783.09 ^{b c}	.146	1029.76 ^d	.454	959.82 ^{d e}	.314
13. Listening (spoken word)	62	55.29	67.49	666.82 ^{a b}	30	575.94 ^{a b}	-.214	786.57 ^c	.092	910.83 ^{d e}	.155

Notes: Identical letters in the superscripts in the median columns mark subsets of activities for which the distribution of the target variables' scores do not significantly differ ($p > .05$). Conversely, different letters mark sets between which significant differences were identified by the post-hoc analyses to the Kruskal-Wallis test.

9.2.4 Relationships Between the Different Aspects of Engagement

As discussed previously, measures of the participants' general levels of affective, cognitive, and linguistic engagement were derived from the Multi-level CFA (Section 9.3.1.5). Multi-level modelling also allowed the researcher to explore how the various dimensions of engagement related to each another at the activity-level (within one activity) and at the participant level (individual tendencies). For this purpose, the researcher added the measures of quantity and variety of ISLPs introduced above to the MCFA as manifest variables. The correlations between the different engagement dimensions are presented in Table 9.14 below.

Table 9.14: Multi-level Confirmatory Factor Analysis of engagement, correlations between different dimensions of engagement at the activity- and participant-levels (r-coefficients)

	<i>Quantity of ISLL</i>	<i>Affective Engagement</i>	<i>Cognitive Engagement</i>	<i>Linguistic Engagement</i>
PARTICIPANT LEVEL				
<i>Quantity of ISLPs</i>	—			
<i>Affective Engagement</i>	.06	—		
<i>Cognitive Engagement</i>	.28 *	.74 ***	—	
<i>Linguistic Engagement</i>	-.01	.24 *	.66 ***	—
<i>Variety of ISLPs</i>	.52 ***	-.01	.10	-.02
ACTIVITY-LEVEL (PER DIARY ENTRY)				
<i>Duration (Quantity)</i>	—			
<i>Affective Engagement</i>	.18 **	—		
<i>Cognitive Engagement</i>	-.12 *	.54 ***	—	
<i>Linguistic Engagement</i>	-.17 **	.19 *	.67 ***	—

Estimation method: Robust Maximum Likelihood with listwise deletion. $\chi^2_{[47]} = 169.08, p < .001$; $n_{obs} = 1505$; $n_{part} = 254$; $CFI = .97$, $TLI = .95$, $RMSEA = .05$; $SRMR_{w/i} = .05$; $SRMR_{b/w} = .05$.

The correlations between the various dimensions of engagement were notably different at the two levels of analysis. At the activity-level, cognitive engagement correlated strongly with affective ($r = .54$) and linguistic engagement ($r = .67$). This means that, during activities on which they were more focused, participants experienced higher levels of interest and enjoyment and also focused more on

linguistic features they encountered. Affective and linguistic engagement, in turn, only shared a weak (but still significant) correlation ($r = .19$), indicating that participants paid slightly more attention to form in activities which they enjoyed. Finally, the activity duration shared a weak positive correlation with affective engagement ($r = .18$) and weak negative correlations with cognitive and linguistic engagement (respectively: $r = -.12$ and $-.17$). This means that participants experienced slightly more interest and enjoyment during activities on which they spent more time but were also less concentrated and did not focus on the language as much.

At the participant-level, similar but stronger correlations were found between cognitive, affective, and linguistic engagement: Cognitive engagement was highly correlated to affective and linguistic engagement (respectively: $r = .74$ and $.66$). This means that participants who were generally more focused on the activities in which they engaged were also likely to report higher levels of enjoyment and focus on the linguistic features. Meanwhile, affective and linguistic engagement were only weakly correlated ($r = .24$), meaning that higher levels of emotional engagement and conscious linguistic engagement co-occurred less often in the same participants. It is worth noting that, at the activity-level, the correlation between cognitive and linguistic engagement was stronger than the link between cognitive and affective engagement, whereas it was the opposite at the participant-level, illustrating why worth to analysing engagement separately at both levels.

When it came to the relationships between the quantity of ISLPs and the other engagement dimensions, the differences between the two levels were even greater: At the participant-level, there were no significant correlations between quantity and affective or linguistic engagement, whereas at the activity-level quantity shared a small positive correlation with affective engagement ($r = .18$) and a small negative

correlation with linguistic engagement ($r = -.17$). Furthermore, at the participant-level, there was a moderate positive correlation between cognitive engagement and quantity ($r = .28$), which had been *negative* at the activity-level. Thus, while participants were generally inclined to focus less during longer activities (negative activity-level correlation), the participants who spent more time on in total on ISLPs also tended to record higher levels of cognitive focus (positive participant-level correlation). Finally, the participant-level variable measuring the number of unique activities in which each participant engaged per day (variety) was strongly correlated with the quantity of ISLPs ($r = .52$), but not with any of the other engagement dimensions. This indicates that the same participants who engaged in a greater variety of activities also tended to spend more time on ISLPs every day.

9.2.5 Are There Distinct Overall Patterns of Engagement in ISLPs Among the Participants?

Since there is often vast individual variation in engagement (confirmed in previous studies [Section 2.1.1] and illustrated by the high within- and between-participant variability in the engagement scores discussed before), researchers have suggested that more person- rather than variable-centred methods should be employed in the study of this construct (e.g. Eccles & Wang, 2012; Janosz, 2012; Marsh, Lüdtke, & Trautwein, 2009; Taboada Barber, Gallagher, Smith, Buehl, & Beck, 2016). In the current study, *Latent Profile Analysis* (LPA) was therefore used (in addition to variable-focused methods such as CFA and SEM) to investigate whether groups of learners could be identified who showed distinct overall patterns of engagement in ISLPs. LPA is a statistical method used to identify subgroups (or *profiles*) of participants within a sample, based on patterns of correlations between a number of

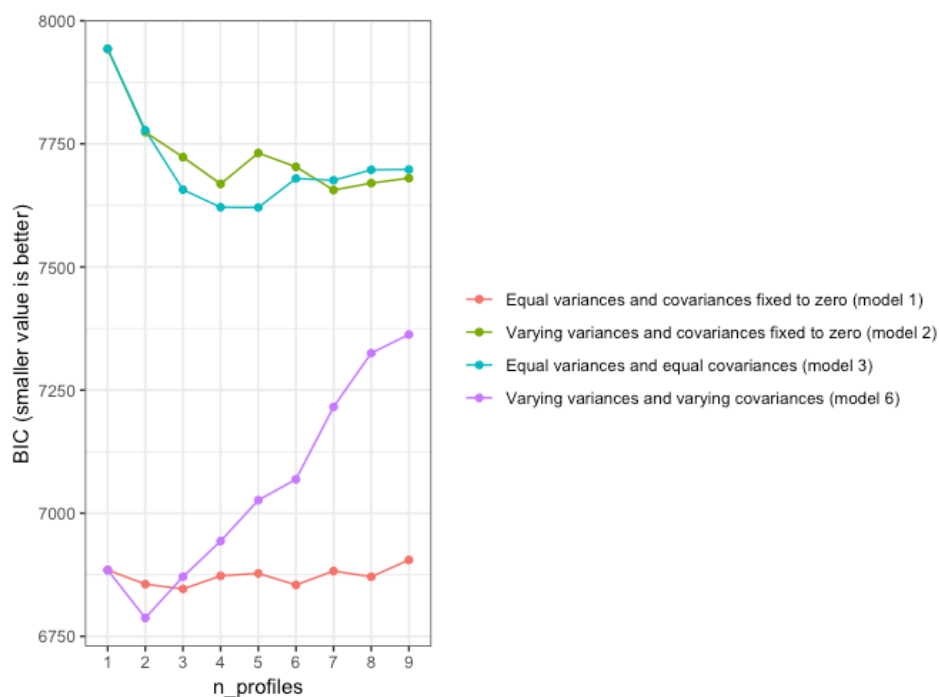
continuous indicators representing an underlying latent factor like engagement (Foorman, Petscher, Stanley, & Truckenmiller, 2017). Thus, LPA is based on the assumption that the way in which several variables interrelate can differ vastly between latent subgroups (G. A. Williams & Kibowski, 2015).

There is no general consensus regarding the minimum sample size required to conduct LPA, since it is dependent on the number of profiles, as well as the number and quality of the indicators (Marsch, Hau, Balla & Grayson, 1998; Finch & Bronk 2011). Simulations have suggested that a sample size of around 200 can be sufficient when there are six good quality indicators (Carlson Wurpts, 2012). In the present study, five indicators were specified in the LPA (the FS for affective, cognitive, and linguistic engagement derived from the MCFA and the measures of quantity and variety of the ISLPs) and complete data were available for 254 students. Thus, the sample was deemed sufficient to conduct the planned analysis. As in the previous CFAs, the class to which the participants belonged was included as a clustering variable, to control for the nested nature of the sample. The LPA was conducted in R using the package tidyLPA.

It must be noted that profiles identified through LPA do not necessarily reflect existing differences within the population. While LPA can be used to explore data patterns, the findings are highly dependent on the specific sample (i.e. it is a 'bottom-up approach') and the researcher ultimately selects how many profiles to extract. The number of profiles which appears to be optimal may, for example, be influenced by non-linearity in the relationships between the indicator variables and factors not included in the model could also be responsible for some of the heterogeneity in the data (Bauer & Curran, 2004). Patterns identified in LPA should therefore not be accepted without further validation.

There are several types of models in LPA, which differ according to which parameters are estimated freely (vs. fixed): In the most constrained model (*Model 1*), the variances of the specified indicators are constrained to be equal across all profiles and the relationships between the variables are not estimated (Rosenberg, 2018; Williams & Kibowski, 2015). In more complex models, on the other hand, these parameters may be allowed to vary to a greater extent, thus adding information by which the characteristics of the profiles can be understood. To identify the best solution, in terms of model type and the number of extracted profiles, the LPA software provides a graphical overview of the *Bayesian Information Criterion* (BIC) for the four most common model types with a varying number of profiles (Figure 9.15). In general, lower BIC values indicate a better model fit, but it is an adjusted goodness-of-fit index, which also penalises more complex models to avoid ‘overfitting’ (Rosenberg, 2018). In this analysis, Model 6 with two profiles clearly emerged as the best solution according to the BIC.

Figure 9.15: BIC values for solutions extracted by Latent Profile Analysis



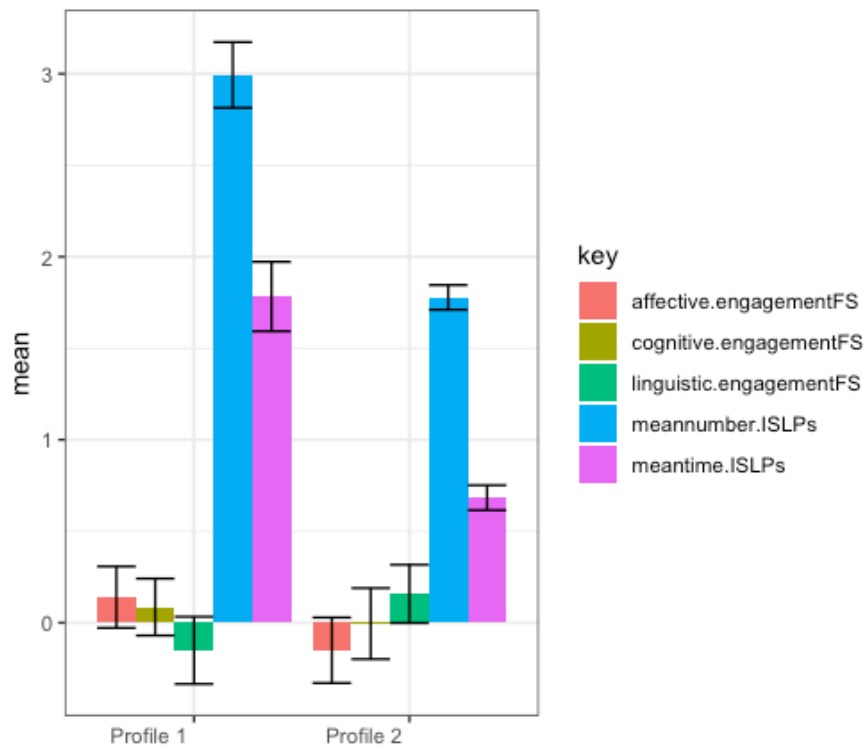
Another way to judge how well a model is reflected within the data is to look at *Entropy*. This is a measure of how distinct the identified groups are in terms of the indicator variables, which is estimated by comparing the variation within profiles to the separation between them (B. Muthén, 2004). Values approaching 1 indicate a clearer delineation of the profiles and .8 is frequently mentioned as a threshold (Celeux & Soromenho, 1996; Ramaswamy, DeSarbo, Reibstein, & Robinson, 1993). In the present study, Model 6 with two profiles returned an entropy value of .94, which reflects an excellent model fit. It was therefore decided to explore this solution in more detail. Both profiles in this model happened to include an approximately equal number of participants ($n_{Grp1} = 126$ and $n_{Grp2} = 128$). Table 9.15 reports the means and standard deviations of the five indicator variables for the two profiles. Figure 9.16 is a graphical comparison of the scaled group means.

Table 9.15: Descriptive statistics of the five indicators for the two groups extracted by LPA.

<i>Variables</i>	<i>Group 1</i> <i>n = 126</i>	<i>Group 2</i> <i>n = 128</i>	<i>Group differences</i>
	<i>Mean Rank</i>	<i>Mean Rank</i>	
<i>Affective Engagement</i>	138.79	116.39	$U = 6642, z = -2.43, p < .05$
<i>Cognitive Engagement</i>	129.52	125.51	$U = 7809, z = -0.44, p = .66$
<i>Linguistic Engagement</i>	114.16	140.63	$U = 6383, z = 2.87, p < .01$
<i>Mean time on ISLPs per day (in minutes)</i>	172.64	83.06	$U = 2376, z = -9.72, p < .001$
<i>Mean number of ISL activities per day</i>	172.91	82.80	$U = 2342, z = -9.98, p < .001$

Note: Grey rows mark variables according to which the two LPA-derived profiled differed significantly. The affective, cognitive, and linguistic engagement values are based on the standardized factor scores derived from the MCFA. That is, they have an overall sample mean of 0 and standard deviation of 1.

Figure 9.16: LPA, comparison of the scaled mean scores for the two extracted groups.



To find out which of these scores were significantly different between the groups, a series of Mann-Whitney U tests were conducted, the results of which are also reported in Table 9.15. This non-parametric alternative to the independent samples t-test was chosen because of the non-normal distribution of the engagement variables within the two groups. These tests indicated that the participants belonging to the first group spent significantly more time on ISLPs than those in Group 2 (approximately 160 vs. 60min/day; $U = 2376, z = -9.72, p < .001$) and also engaged with a greater variety of activities per day (approximately 1.6 vs. 1; $U = 2342, z = -9.98, p < .001$). Group 1 also recorded significantly higher average levels of affective engagement ($U = 6642, z = -2.43, p < .05$), although the opposite was true for linguistic engagement ($U = 6383, z = 2.87, p < .01$). When interpreting these results it must be kept in mind that affective and linguistic engagement were operationalised as standardised factor scores: With zero representing the sample-wide mean, the CFA

group means thus indicated below-average linguistic engagement and above-average affective engagement in Group 1, and vice versa in Group 2. In terms of cognitive engagement, there was no significant difference between the two groups ($U = 7809, z = -0.44, p = .66$).

9.2.5.1 Engagement Profiles and Types of Informal Activities

A series of follow-up analyses were conducted to find out more about the two profiles extracted from the LPA. In order to make modelling easier, group membership was treated as a manifest variable during these analyses. That is, the analyses were based on the assumption that the LPA software accurately assigned each participant to the correct group. However, the group membership of each participant is predicted by the latent profile model on the basis of probabilities, that is, there is some uncertainty (error variance). The tidyLPA software expresses this uncertainty as a posterior probability, which represents the estimated likelihood that the model has assigned a participant to the correct group. In the present analysis, these posterior probabilities were very high ($M = .94, SD = .10$), supporting the assumption that the extracted groups represent a sufficiently reliable categorisation of the participants based on the engagement variables entered into the LPA model.

Follow-up tests were conducted to examine whether there was a difference between the two LPA groups in terms of the types of activities in which they engaged and how long they spent on these activities. As with gender (Section 9.2.2.1), this was done through a series of logistic and simple linear regressions. The results of these analyses are reported in Appendix F.5.

The engagement profile significantly predicted whether or not participants engaged in mixed social media ($p < .001$) but not in any of the other activities. The participants who in LPA Group 1 were more than 3 times as likely as the participants in Group 2 to use mixed social media and group membership explained 11% of the variance in this variable (Nagelkerke r^2 , small to medium-sized effect). Meanwhile, the linear regression analyses revealed that participants in Group 1 also spent significantly more time (in decreasing order of effect size) playing video games while chatting with other players, watching serials, listening to spoken English (e.g. audiobooks, conversations), using mixed social media, listening to music (both in the fore- and background), and using video-based social media. Altogether, the LPA engagement profile had small to medium sized effects on the time spent doing these activities (r^2 between .04 and .18).

9.2.6 Key Findings

The first research question asked how the pupils in this study engaged in ISLL. There was considerable variation in the amount of time that individual participants spent on informal second language practices: On average, they dedicated about 1.8 hours per day to ISLPs ($M = 108.3\text{min}$, $SD = 87.6$), but overall, the values ranged from only four minutes per day to as high as 9 hours (539min). The individual participants' informal activities were not particularly varied, with most students engaging, on average, in only one to two unique activities per day. Listening to music, using social media, and watching serials were the three most popular activities in this sample, in terms of the frequency of engagement, the time spent on these activities, and the proportion of students who engaged in them. The least popular activities (by the same criteria) were listening to spoken English, watching films, and writing. Video gaming

stood out as an activity which ranked comparably low in terms of overall popularity, but on which by far the longest *continuous* amount of time were spent ($M = 170\text{min}$, $SD = 111.7$). More specifically, gaming while also communicating with other players was among the activities during which the participants experienced the highest levels of affective and cognitive engagement but ranked among the lowest in terms of linguistic engagement. Of the other activities, those that focused more explicitly on language processing or production (reading books, writing, and speaking in English) seemed to be associated with higher levels of cognitive and linguistic engagement and lower levels of affective engagement than those aimed more at entertainment (e.g. using social media, watching serials).

Looking at the correlations between the different types of engagement at the activity-level (per diary entry) revealed that there were significant relationships between all engagement measures. Generally, these were weak relationships ($r < |.20|$), except for the correlations of cognitive engagement with affective and linguistic engagement (both $r > .50$). This suggests that pupils generally experienced higher levels of enjoyment and interest and also paid more conscious attention to linguistic features during activities on which they were more focused. Similarly, strong relationships between cognitive engagement on the one hand and affective and linguistic engagement on the other were also found at the participant-level, indicating that pupils who generally had a greater tendency towards focusing on the informal activities themselves (vs. extraneous stimuli) also experienced more enjoyment and focused more closely on linguistic features they encountered. At the participant-level, there was also a strong relationship between the two measures of behavioural engagement (quantity and variety), indicating that students who spent more time on ISLPs also tended to engage in a varied repertoire of informal activities. The only

other significant correlations at this level were between cognitive engagement and the quantity of ISLPs, and between affective and linguistic engagement (both $r \approx .25$).

Finally, a Latent Profile Analysis was conducted to investigate whether the overall engagement patterns varied across different subgroups within the sample. Two groups were extracted, each of which contained an approximately equal number of participants. The pupils in Group 1 spent significantly more time on ISLPs per day ($M = 160.1\text{min}$, $SD = 96.9$ vs. $M = 60.4$, $SD = 34.3$) and their practices were significantly more varied (unique activities per day: $M = 1.62$, $SD = .57$ vs. $M = .97$, $SD = .22$). The groups also differed significantly with regard to affective and linguistic engagement, with the participants in Group 1 recording higher levels of enjoyment and interest in ISLPs, but lower levels of linguistic engagement. The groups did not differ significantly with regard to cognitive engagement.

In Section 9.2 the researcher described the engagement of this study's participants in terms of various different aspects (behavioural, affective, cognitive, and linguistic). Sections 9.3–9.5 are dedicated to exploring the relationships between these engagement dimensions and other variables: proficiency, motivation, and attitudes towards language learning.

9.3 Research Question 2: What is the Direct Relationship Between General Language Proficiency and Engagement in Informal Second Language Practices?

The analyses presented in this section were conducted to investigate the direct relationships between developing general language proficiency (as measured by the C-test scores at two time points, three months apart) and engagement in informal second language practices (ISLPs). Three sets of analyses were carried out, which

investigated the relationships of general language proficiency with (1) engaging in different types of informal second language activities (Section 9.3.1); (2) the various dimensions of engagement: behavioural, affective, cognitive, and linguistic engagement (Section 9.3.3); and (3) the groups discovered in the LPA analysis, which represent the participants' general patterns engagement in ISLPs (Section 9.3.4).

9.3.1 Proficiency and Types of Informal Second Language Activities

Previous ISLL studies have primarily tried to find links between the time spent on informal second language practices and language proficiency (see Section 2.1.2). However, some studies have also looked at whether particular informal activities are more strongly correlated with proficiency than others. In order to confirm and expand these findings, the researcher decided to include a set of analyses to explore whether (a) participants' prior proficiency could predict the types of activities with which they chose to engage, and (b) whether the development of proficiency over the time of the study could be predicted by their engagement in different activities. As in the previous analyses concerning whether gender and LPA group membership could predict engagement in the different the types of informal activities (Sections 9.2.1.1 and 9.2.5.1), a multi-step analytical approach was adopted to deal with the zero-inflated distributions of the variables representing the time spent on specific activities. The results are reported in full in Table G.1 and G.2, Appendix G.1 and summarised below.

The significant relationships between prior proficiency and engagement in different activities indicated in the logistic regressions were positive in all cases except for the variables coding for playing video games, both with and without concurrent chatting to other players. This means that, on the whole, more proficient students were more

likely to engage most of the informal activities. However, these regression paths reached statistical significance in only a few instances, namely (in order of effect size, from strongest to weakest): watching films or serials; listening to spoken English; using mixed social media; writing; focused music listening; and video-based social media use (all $p < .05$, r_N^2 between .03 and .07, indicating small effects). It is worth noting that this list includes almost all activities which mostly involve listening to English (except for playing music in the background). A slightly different picture emerged from the simple regressions, where prior proficiency only significantly predicted the time spent on social media (all three types) and reading (all $p < .05$, r^2 between .06 and .14, indicating small to medium effects).

The second question (whether engagement in certain activities could predict changes in proficiency over the course of this study) was addressed through two stepwise regression analyses, with proficiency at T1 entered as a predictor in the first step, followed by the two sets of variables coding for activity type (binary and continuous) as predictors of proficiency at T2. In this analysis, a vast majority of the variance in proficiency at T2 was explained by the prior proficiency measure ($\beta = .98$, $p < .001$, $r^2 = .87$). In fact, none of the individual informal activities (whether or not participants engaged in them at all or how much time they spent on them per day) predicted a significant amount of the variance in proficiency at T2 not already by prior proficiency.

The findings presented in the remainder of Section 9.3 focus on the relationships between proficiency and (1) the different dimensions of engagement in ISLPs and (2) the general engagement patterns exhibited by the two groups extracted through

LPA. The main statistical approaches for these analyses were *Structural Equation Modelling* and *Path Modelling*. Before presenting the findings that emerged from these analyses (Sections 9.3.3 and 9.3.4), these methods are briefly introduced in the next section.

9.3.2 Structural Equation Models and Path Modelling

There are two components to Structural Equation Models (SEMs): the *Path Model* and the *Measurement Model*. Similar to regression analyses, Path Models use variances and covariances to investigate hypothesised linear relationships between predictor and outcome variables. However, whereas regression analyses are restricted to studying the *direct* effects of one or multiple predictors, path models are capable of testing more complex relationships between variables (e.g. Blunch, 2013; R. B. Kline, 2015; Tabachnik & Fidell, 2014). This includes mediation models, which seek to explain the observed relationship between variables via the *indirect effect* of one or more intermediary constructs. In the present study, path models then provide a way to explore whether different dimensions of engagement in ISLPs can predict, as mediating variables between proficiency, motivation, and attitudes measured at two time points, some of the variation in these variables over the time of this study.

The *Measurement Model* comes into play when the variables under investigation are 'latent', that is not directly measurable (e.g. psychological constructs). Using CFA, this part of the SEM tests whether hypothesised latent factors are well represented in the collected data. Thus, in contrast to regression, SEMs can take into account the measurement quality where instruments such as multi-item scales are used as indirect indicators of latent constructs.

Traditionally, this is done by formulating and testing a *Latent Variable Structural Model*, which combines both the measurement and the path model into one, estimating all factor loadings at the same time as the variances and hypothesised covariances between the constructs. The full Latent Variable Structural Model required to answer RQ2 in this study, which places the five engagement measures as mediators between general language proficiency at T1 and T2, is pictured in Figure G.1. Appendix G.2. As in the previous Multi-level Confirmatory Factor Analyses (Section 9.1.3.5), engagement would need to be modelled separately at the within- and between-participant levels. As the schema in Appendix G.2 makes clear, this model is relatively complex, including 33 manifest indicators and 8 latent factors and requiring the estimation of 105 total parameters (32 factor loadings, 44 variances, 18 covariances, and 11 regression paths).

The complexity of the model which is to be tested, along with the chosen estimation method and the size of the latent factor loadings and path coefficient that are to be detected, raises issues around the sample size required to conduct an SEM analysis (L. K. Muthén & Muthén, 2002; Wolf, Harrington, Clark, Shaunna, & Miller, 2013). If sample sizes are too small, error rates can be inflated, that is the accuracy of the estimated parameters may suffer. There are many rules of thumb for the appropriate sample size in SEM, which focus on the ratio of the sample size to the number of estimated parameters. Bentler and Chou (1987), for example, suggested that as little as 5 observations per parameter could be sufficient, although most researchers prefer a ratio of at least 10 or 20 (e.g. Hoogland & Boomsma, 1998; R. B. Kline, 2015). With a ratio of 10.6 observations per estimated parameter, the available sample size was judged by the researcher as sufficient for analysing the model of engagement with

proficiency as pictured in Figure G.1, Appendix G.2. However, some of the other planned models (discussed in Sections 9.4 and 9.5) included an even greater number of variables and were thus judged to be too complex to be reliability tested with the available data.

One of the strategies that has been suggested to deal with sample size limitations is to analyse the measurement model (CFA) separately and use it to derive latent variable factor scores (FS), which can then be used in path modelling (Devlinger & Rosseel, 2017). This *Factor Score Path Modelling* technique was adopted in the later analyses in this study, using the factor scores generated from the earlier CFAs (Section 9.1.3).

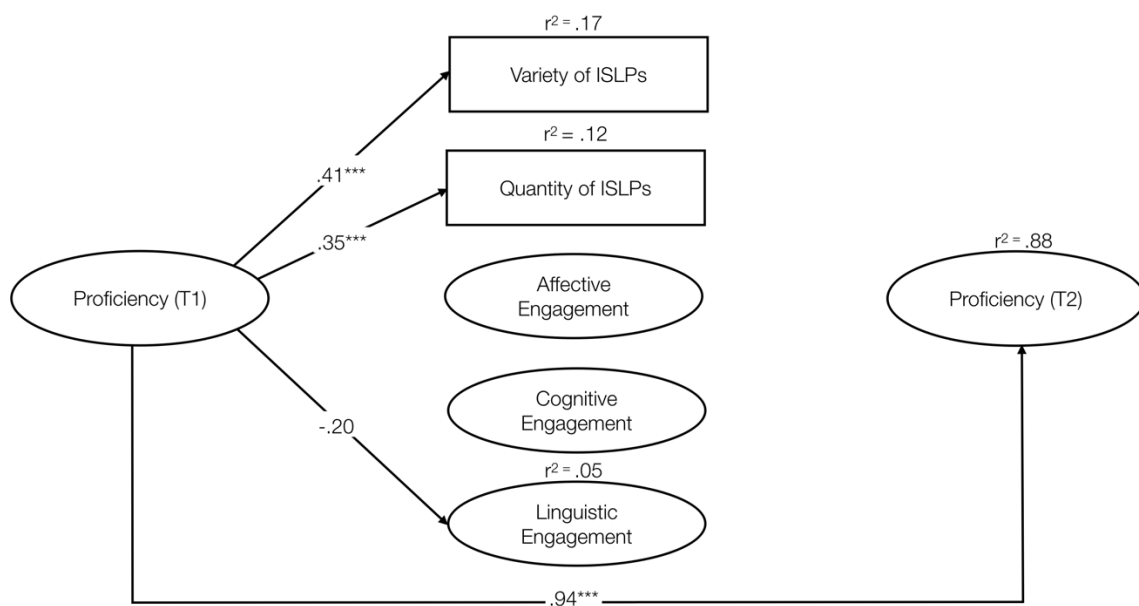
The following section reports the outcomes of the Multi-level Structural Equation Model of proficiency and the five measures of engagement in ISLPs. Additionally, the section also summarises the results of a series of comparative analyses which were done to investigate whether the quantitative findings were influenced by the missing engagement data patterns and to confirm whether the Factor Score Path Modelling approach used in the subsequent analyses would deliver results similar to those of a true Multi-Level Latent-Variable SEM analyses (Section 9.3.3).

9.3.3 Proficiency and the Different Dimensions of Engagement

The first SEM focused on the relationships between the five measures of engagement in ISLPs and proficiency, measured across two time points, three months apart (full model depicted in Figure G.1, Appendix G.2). As in the preceding CFAs, Robust Maximum Likelihood Estimation was used to obtain goodness-of-fit indices which are robust against non-normality. Missing data were deleted listwise, as the lavaan

software currently does not provide the option imputing missing values through Full Information Maximum Likelihood Estimation for multilevel models. Overall, the fit indices (reported at the bottom of Figure 9.17) suggested a good model fit. Figure 9.17 shows the significant regressions at the participant-level, including the standardised regression coefficients (β -values) and the amount of variance in each outcome variable cumulatively explained by the included predictors (r^2 -values). The unstandardized coefficients and standard errors are can be found in Table G.3, Appendix G.2.

Figure 9.17: Proficiency and engagement Multi-Level Structural Equation Model. Significant regression paths at the participant-level



Estimation method: Robust Maximum Likelihood with listwise exclusion. $\chi^2_{[244]} = 442.24$, $p < .001$; $n_{obs} = 1113$, $n_{part} = 188$; $CFI = .96$, $TLI = .95$, $RMSEA = .03$, $SRMR_{w/i} = .04$, $SRMR_{b/w} = .07$.

None of the five engagement variables significantly predicted changes in the proficiency scores from T1 to T2. The only significant predictor of posterior proficiency was proficiency at T1, which explained 88% of the variance in the T2 scores.

In fact, affective and cognitive engagement showed no relationship to proficiency at all. The two measures of behavioural engagement, the variety and quantity of ISLPs, were both predicted significantly by proficiency at T1 ($\beta = .41$ and $.35$, both $p < .001$), with r^2 values around $.15$ indicating small to medium-sized effects. Meanwhile, the path from prior proficiency to linguistic engagement was only marginally non-significant ($\beta = -.20$, $p = .052$), suggesting that there may be a very weak negative relationship between the two variables.

A closer look at changes in proficiency across the study

Since a very large proportion of the variance in the proficiency scores at T2 seemed to be explained by the T1 scores (88%), the researcher decided to explore whether the C-test did actually capture any changes in proficiency over the course of the study. For this purpose, sum scores were calculated for each participants' proficiency at T1 and T2, which were weighted according to the factor loadings in the C-test measurement invariance models.

A paired-samples t-test revealed that there was indeed a significant difference between these the C-test outcomes at both time points ($t[244] = -10.59$, $p < .001$, $d = .46$, indicating a small to medium-sized effect), with scores increasing, on average, by 7.3% (from $M_{T1} = 51.40\%$, to $M_{T2} = 58.69\%$; $SD_{T1} = 18.24\%$ and $SD_{T2} = 19.84\%$). This tentatively suggests that there was indeed an increase in the proficiency of the participants over the course of this study. However,

it must be kept in mind that these mean scores should only be taken as a rough indication, since the two C-test versions did not achieve strict measurement invariance (see Section 9.1.3.1).

The effects of missing engagement data

As with the Multi-Level CFA (Section 9.1.3.5), two additional analyses were run to test the extent to which the above findings might have been biased by data from participants who submitted diary entries across only a few days. The Multi-Level SEM model was tested twice more, excluding observations from participants who made entries over fewer than four or five days (respectively: $n_{part} = 45, n_{obs} = 70$ and $n_{part} = 62, n_{obs} = 150$). It should be noted that, while these exclusions may improve the accuracy of the participant-level parameter estimates, there is a trade-off with the power (and therefore overall accuracy) of the analysis, as these exclusions reduce the ratio of observations to parameters to 9.93 and 9.08 respectively (i.e. *just* under the minimum sample size requirements, see Section 9.3.2). The full results of these analyses are presented in Figures G.2 to G.4, Appendix G.2 and summarised below.

The exclusions did not significantly affect the overall model fit or the variable covariances. The variety and quantity of ISLPs were still significantly predicted by T1 proficiency. In addition, the previously only marginally significant path from T1 proficiency to linguistic engagement now also reached statistical significance, although the effect size remained small (Model 2: $\beta = -.22, p = .050, r^2 = .05$; Model 3: $\beta = -.23, p = .018, r^2 = .08$). There were minor changes to the magnitude, and sometimes directionality, of some of the other regression coefficients, but these can be disregarded as they did not reach significance in any of the three models.

Comparison of SEM and factor score path modelling

To examine the effect of using Factor Score Path Modelling to address the subsequent RQs (Sections 9.4 and 9.5) instead of directly modelling the latent variables using Multi-level SEMs, the researcher also tested the same model as above using the CFA-derived factor scores for proficiency, affective, cognitive, and linguistic engagement. The outcomes are reported in full in Figure G.5, Appendix G.2 and summarised below.

In essence, this FS path model delivered the same results as the Multi-level SEM presented above: Only the paths from proficiency at T1 to proficiency T2, the quantity, and variety of engagement reached significance. The path towards linguistic engagement remained marginally significant ($\beta = -.18, p = .056, r^2 = .04$). The significant standardised regression coefficients and effect sizes were also approximately within the same order of magnitude as those derived from the multi-level model, thus leading to the same conclusions about the comparative strength of these relationships. There were some differences between the two models in the magnitude, and sometimes directionality, of some of the other regression coefficients, but once again, these can be disregarded as they did not come close to reaching statistical significance in either model.

9.3.4 Proficiency and Engagement Profiles

To explore the relationship between the participants' general patterns of engagement in ISLPs (the groups discovered in the LPA, Section 9.2.5), the researcher planned to test a simple single-level SEM in which a binary *engagement profile* dummy variable (representing the LPA groups; Group 1 = 0; Group 2 = 1) would be entered as a mediator between proficiency at T1 and T2 (Figure 9.18). However, this model could not be brought to converge, perhaps reflecting a sample size issue: For this analysis

the ratio of data points to parameters was only 3.8, that is, far below the commonly accepted critical threshold of 10. Even more, larger sample sizes are generally required for SEMs such as this, which involve categorical instead of continuous variables and therefore must use *Diagonally Weighted Least Squares* instead of Maximum Likelihood Estimation (Bandalos, 2014; Forero, Maydeu-Olivares, & Gallardo-Pujol, 2009).

To reduce the number of free parameters (thus increasing the ratio of data points to parameters), the researcher decided instead to conduct the same analysis using a Factor Score Path Model (Figure 9.19).

Figure 9.18: Proficiency and engagement profile, planned Structural Equation Model

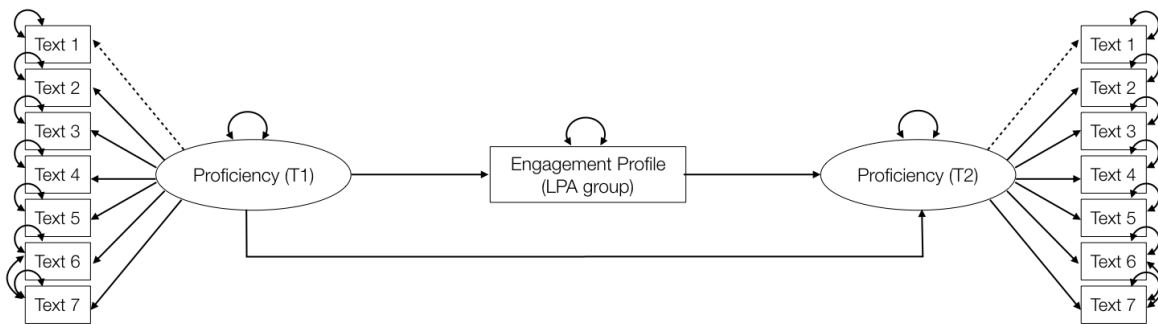
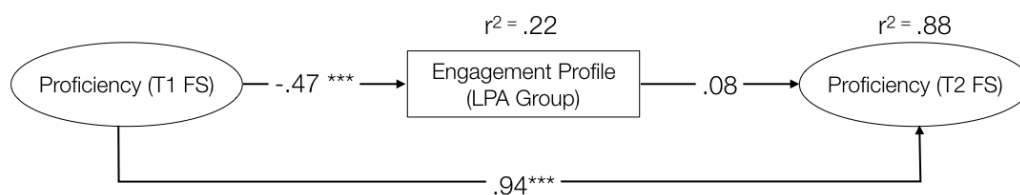


Figure 9.19: Proficiency and engagement profile, Factor Score Path Model



Estimation method: Diagonal Weighted Least Squares with pairwise exclusion.

Note: Fit indices are not available because this is a saturated model (0 df).

The LPA groups, which represented participants' general patterns of engagement, did not, in fact, significantly predict any changes in the participants' proficiency over the course of this study ($\beta = .08, p = .840$). Group membership was significantly

predicted, however, by the prior proficiency scores ($\beta = -.47, p < .001$). The negative coefficient indicates that participants who had higher proficiency at T1 were more likely to be part of Group 1, which showed significantly higher levels of affective and behavioural engagement and lower levels of linguistic engagement. Prior proficiency explained 22% of the variance in the binary engagement profile variable, indicating a medium to large effect size.

9.3.5 Key Findings

The analyses presented in this chapter investigated the direct relationships between general language proficiency and engagement in informal second language practices. Three sets of analyses were carried out, in each which engagement was operationalised in different ways: (1) through variables which represented whether or not participants engaged in certain types of informal activities; (2) through five variables measuring various dimensions of engagement (quantity and variety of informal practices [behavioural engagement], affective, cognitive, and linguistic engagement); and (3) through a binary *engagement profiles* variable coding for the groups extracted via LPA in the previous chapter, which operationalised the participants' general patterns of engagement in ISLPs.

Contrary to what had been hypothesised (based on the wider ISLL literature, see Section 2.2), engagement did not predict changes in proficiency between T1 and T2 in any of the analyses. However, in all of them, prior proficiency did significantly predict some of the aspects of engagement in informal second language practices.

The first set of analyses looked at the relationship between proficiency and engagement in different types of informal activities. Only in a few cases did prior

proficiency significantly predict how likely participants were to engage (or not) in certain informal practices. These were: watching films, watching serials, listening to spoken English, mixed social media use (i.e. involving both video- and text-based activities), listening to music as a focused activity, and purely video-based social media use (all R_N^2 between .03–.07, indicating small effects). It is noteworthy that this list includes almost all activities which involve the listening skill, with the exception of playing music in the background. Furthermore, proficiency at T1 predicted the amount of time spent on only four of the thirteen activities: all three types of social media use (text-based, video-based, and mixed) and reading (r^2 between .06 and .14, indicating small to medium-sized effects).

In the second analysis, proficiency at T1 significantly predicted the quantity and variety of informal practices in which the participants engaged (respectively, $r^2 = .12$ and $.17$, indicating medium-sized effects), suggesting that more proficient students spent, on average, more time on ISLPs and engaged in a wider range of informal activities. Furthermore, there was a negative relationship between prior proficiency and linguistic engagement which was only marginally non-significant ($p = .052$), implying that students with lower proficiency scores at T1 were more inclined to focus consciously on the linguistic features they encountered in their informal second language practices.

Finally, proficiency at T1 was also a significant predictor of membership in one of the two LPA groups, which represented the participants' general patterns of engagement in ISLPs. Specifically, participants with higher prior were more likely to be part of LPA Group 1, whose members showed significantly higher average levels of affective and behavioural engagement (both quantity and variety of ISLPs) and lower levels of linguistic engagement than those in Group 2.

Section 9.3 addressed the direct relationships between engagement in informal second language practices and general language proficiency. Section 9.4 will focus on the direct relationships between engagement and language learning motivation, as well as attitudes towards formal and informal language learning.

9.4 Research Question 3: What are the Direct Relationships between Language Learning Motivation and Attitudes Towards Language Learning and Engagement in Informal Second Language Practices?

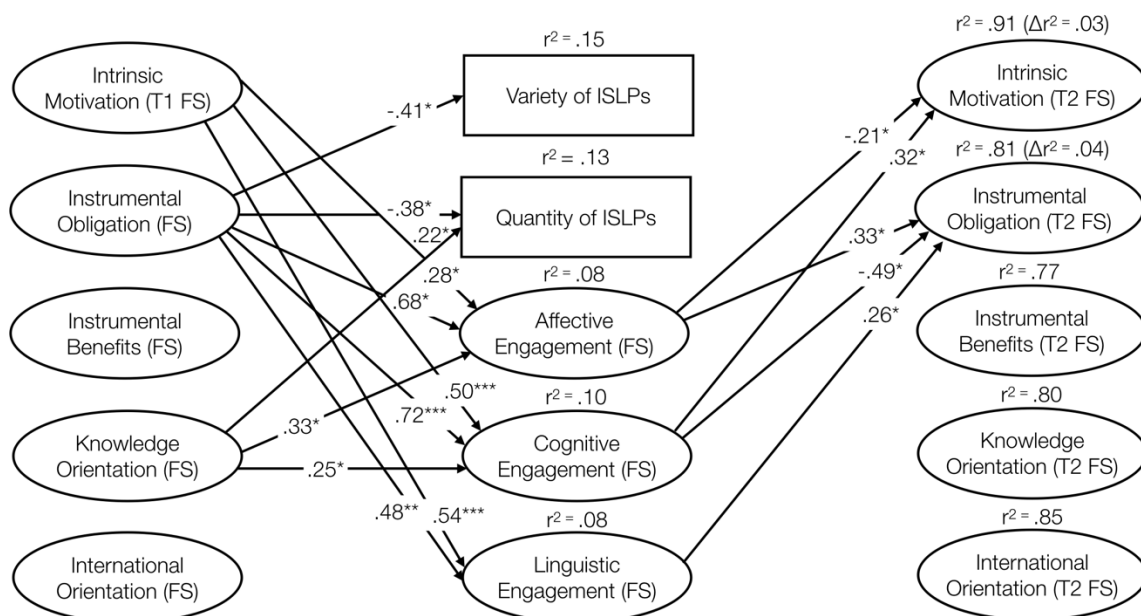
Similar Factor Score Path Models to those presented in the previous section were used to address the second RQ, which was concerned with the direct relationships between engagement in ISLPs and motivation/attitudes. Separate models were formulated to investigate the relationships of the (1) language learning orientations, (2) second language motivational selves, and (3) attitudes towards informal and formal language learning at T1 and T2 with (a) the five measures of engagement in ISLPs (Section 9.4.1) and (b) the LPA groups, coding for general engagement patterns (Section 9.4.2). However, unlike in the previous section on proficiency, the links between motivation and attitudes and the specific types of informal activities in which the participants engaged were not tested, as doing so would have required conducting dozens of separate regression analyses, going beyond what was possible within the scope of this thesis.

9.4.1 Motivation, Attitudes and Dimensions of Engagement

9.4.1.1 Language Learning Orientations

In this model, the five engagement variables were inserted as mediators between the language learning orientations (LLOs) measured at T1 and T2. The LLOs within each time point were also allowed to correlate with one another.¹⁷ This model was an excellent fit for the data (see goodness-of-fit indices reported at the bottom of Figure 9.20). The regression paths which emerged from this analysis as significant (β -coefficients) are shown in Figure 9.20, as well as the total amount of variance explained in each outcome variable (r^2 -values). The non-standardised coefficients and associated p -values are reported in Table H.1, Appendix H.1.

Figure 9.20: Language learning orientations and engagement, Factor Score Path Model. Significant standardised regression path coefficients



Note: Not pictured are the significant paths from the LLOs at T1 to T2. Estimation method: Full Information Robust Maximum Likelihood. $\chi^2_{[7]} = 20.83, p < .01; n = 349; CFI = .99, TLI = .97, SRMR = .01.$

¹⁷ With the exception of instrumental obligation, which showed strong *negative* correlations with all other LLOs, all language learning orientations were strongly positively correlated with one another.

Only three of the five LLOs were significant predictors of engagement in ISLPs. Intrinsic motivation positively predicted (in order of magnitude) the linguistic ($\beta = .54, p < .001$), cognitive ($\beta = .50, p < .001$), and affective engagement scores ($\beta = .28, p < .05$). Second, a strong knowledge orientation was found to predict higher levels of affective ($\beta = .33, p < .05$) and cognitive engagement ($\beta = .25, p < .05$), as well as a more time spent on ISLPs ($\beta = .22, p < .05$). Third, a positive effect of instrumental obligation on cognitive ($\beta = .72, p < .001$), affective ($\beta = .68, p < .05$), and linguistic engagement ($\beta = .48, p < .01$) was found. However, this orientation also had a negative effect on the quantity ($\beta = -.38, p < .05$) and variety of ISLPs ($\beta = -.41, p < .05$). Altogether, the LLOs explained between 8 and 15% of the variance in the different engagement variables, indicating medium-sized effects of the learners' personal reasons for language learning on engagement in ISLPs.

Altogether, the language learning orientations seemed to have a greater effect on engagement in ISLPs than vice versa. Nevertheless, affective, cognitive, and linguistic engagement were found to have significant effects on changes in intrinsic motivation and instrumental orientation over the course of the study (T2 scores, controlling for T1). Curiously, stronger affective engagement significantly predicted a *decrease* in intrinsic motivation ($\beta = -.21, p < .05$), whereas cognitive engagement was associated with an *increase* in this variable ($\beta = .32, p < .05$). The pattern was opposite for instrumental obligation, with higher affective engagement predicting an increase ($\beta = .33, p < .05$) and cognitive engagement predicting a decrease in this variable ($\beta = -.49, p < .05$). These inverse effects are perhaps unsurprising given that intrinsic and instrumental obligation also shared very strong negative correlations at T1 and T2 ($r_{T1}^2 = -.88; r_{T2}^2 = -.92$). Finally, changes in instrumental obligation were also predicted by linguistic engagement ($\beta = .26, p < .05$). This was

a positive link, with greater levels of linguistic engagement predicting an *increase* in instrumental obligation over time. Altogether, affective, cognitive, and linguistic engagement only had a small effect on intrinsic motivation and instrumental obligation, explaining 3–4% more of their variance beyond what was already accounted for by the LLOs at T1.

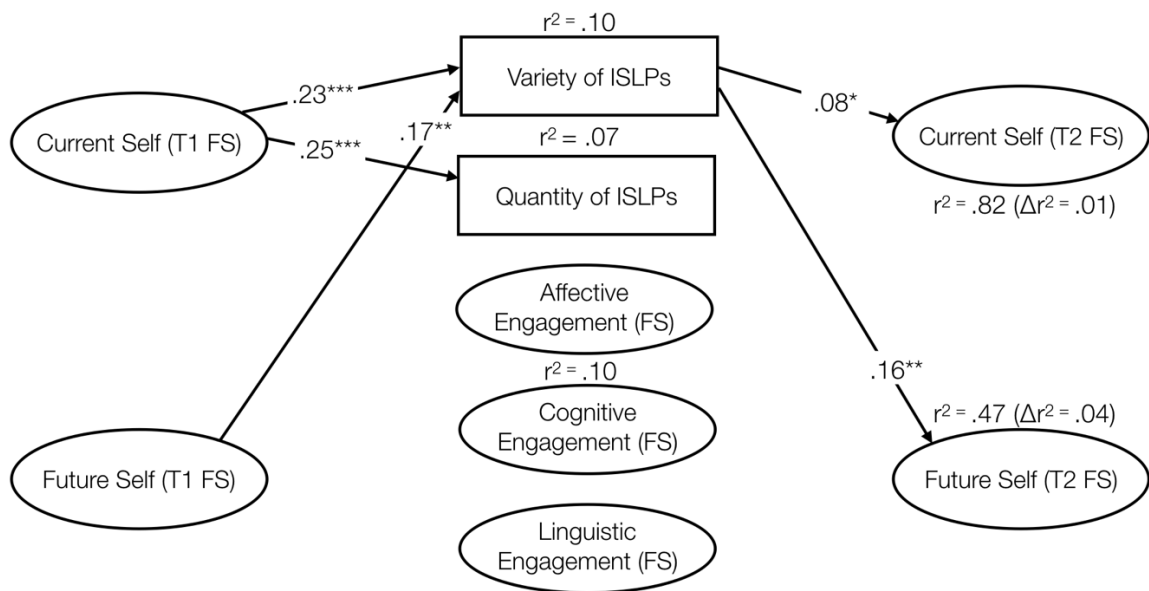
9.4.1.2 Second Language Motivational Selves

Next, the engagement variables were inserted as mediators between the current and future L2 selves at T1 and T2. The variables in each of the three *groups of constructs* (L2 motivational selves at T1, engagement, and motivational selves at T2) were allowed to intercorrelate.¹⁸ The fit indices (reported at the bottom of Figure 9.21) demonstrated that this model was an excellent fit for this data, even though the results in Figure 9.21 and Table H.2, Appendix H.1 show that only a few of the regression paths reached statistical significance. The current self at T1 significantly affected the two measures of behavioural engagement: the time participants spent on ISLPs ($\beta = .25, p < .001$) and the variety of activities with which they engaged ($\beta = .23, p < .001$). Variety of ISLPs was also significantly predicted by the participants' future self-concept ($\beta = .17, p < .01$). Altogether, the current and future L2 selves had small to medium-sized effects on the two measures of behavioural engagement, explaining 7% of the variance in the quantity and 10% of the variance in the variety of ISLPs.

¹⁸ The current and future selves shared significant positive correlations at T1 and T2 ($r_{T1}^2 = .36; r_{T2}^2 = .29$), indicating that participants who wished more strongly to become competent L2 users in future also tended to have a more positive view of their current language proficiency.

As with the LLOs, there were fewer significant effects on the right-hand side of the model, indicating that the motivational selves affected engagement more than vice versa. The variety of ISLPs was the only dimension of engagement which significantly predicted changes in the motivational selves over the course of this study. Variety predicted very small but in the strength of both the current ($\beta = .08, p < .05$) and future L2 selves ($\beta = .16, p < .01$), respectively explaining 1% and 4% in their variance (beyond that already accounted for by the L2 selves at T1).

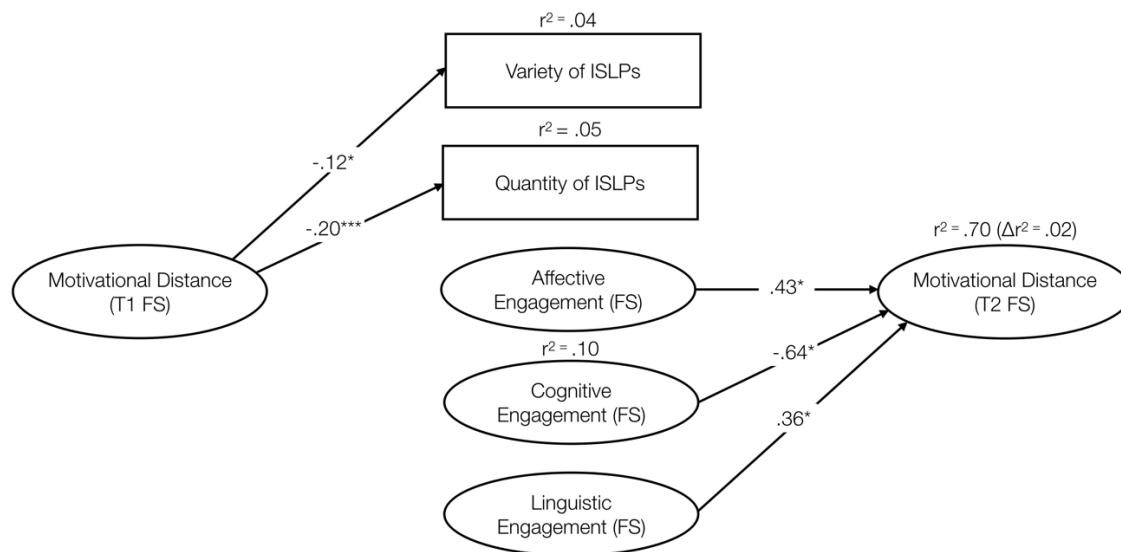
Figure 9.21: Second language motivational selves (current and ideal future self) and engagement, Factor Score Path Model. Significant standardised regression path coefficients



Note: Not pictured are the significant paths from the motivational L2 selves at T1 to T2. Estimation method: Full Information Robust Maximum Likelihood Estimation. $\chi^2_{[2]} = 3.95$, $p = .139$; $n = 349$; $CFI = .99$, $TLI = .97$, $SRMR = .01$.

A separate analysis was thus conducted to test the relationships between the *motivational distance* at T1 and T2 and engagement in ISLPs (Figure 9.22 and Table H.3, Appendix H.1).

Figure 9.22: Motivational distance (difference between current and future self) and engagement, Factor Score Path Model. Significant standardised regression path coefficients



Note: Not pictured are the significant paths from the T1 motivational distance to T2. Estimation method: Full Information Robust Maximum Likelihood. No fit indices available because this model is saturated (0 df).

Like the future and current L2 self, motivational distance significantly predicted variations in the quantity ($\beta = -.12, p < .05$) and variety of ISLPs ($\beta = -.20, p < .001$), although it explained slightly less variance overall (5% and 4% respectively, as compared to the 7% and 10% in the previous model). The regression coefficients were negative, meaning pupils with a smaller motivational distance (more similar current and future selves) spent more time on more varied informal activities.

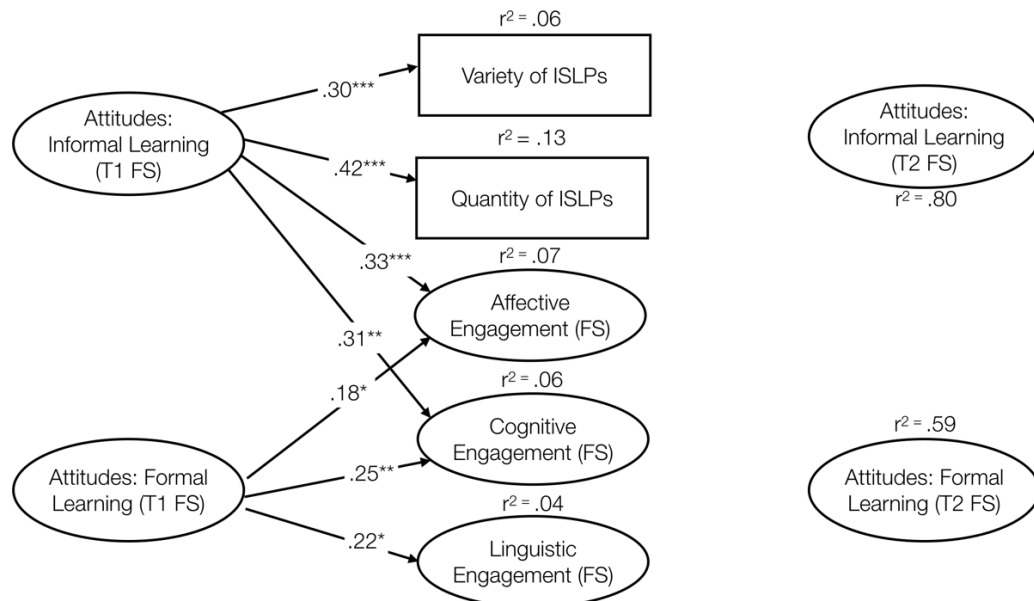
Interestingly, whereas the variety in ISLPs predicted changes in both the current and future L2 self over the course of this study, it did not have a significant effect on the motivational distance. Instead, higher affective and linguistic engagement were both found to predict increases in the motivational distance at T2 (respectively: $\beta = .33, p < .05$ and $\beta = .36, p < .05$), whereas cognitive engagement was associated with a decrease in the motivational distance (i.e. a convergence of the current and future self; $\beta = -.64, p < .05$). Nevertheless, the overall effect of

engagement on the motivational distance was small, with affective, cognitive, and linguistic engagement together accounting for only 2% more of the variance in the motivational distance than what was already explained by the distance between the current and future L2 selves at T1.

9.4.1.3 Attitudes Towards Informal and Formal Language Learning

The participants' attitudes towards language learning were operationalised as their perceptions of the authenticity and effectiveness of learning in informal and formal contexts. The engagement variables were inserted as mediators between the attitude measures at T1 and T2,¹⁹ and each group of variables regressed upon the preceding ones (Figure 9.23 and Table H.4, Appendix H.1).

Figure 9.23: Attitudes towards second language learning and engagement, Factor Score Path Model. Significant standardised regression path coefficients



Note: Not pictured are the significant paths from the attitudes at T1 to T2. Estimation method: Full Information Robust Maximum Likelihood. $\chi^2_{[2]} = 3.67, p = .158; n = 349; CFI = .98, TLI = .94, SRMR = .03$.

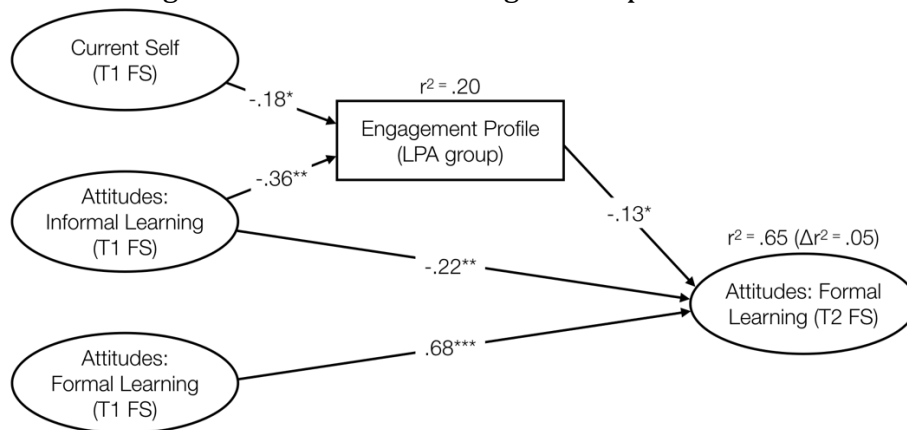
¹⁹ The attitudes were negatively correlated at both time points ($r_{T1}^2 = -.63; r_{T2}^2 = -.28$), indicating that participants who held were more positive about classroom learning tended to believe less in informal learning and vice versa. The correlation was substantially stronger at T1 than T2, suggesting that changes in opposite directions occurred in both variables over the course of this study, which were not predicted by the engagement variables.

Positive attitudes towards informal learning significantly and positively predicted the quantity ($\beta = .42, p < .001$) and variety of ISLPs ($\beta = .30, p < .001$) and the levels of affective ($\beta = .33, p < .001$) and cognitive engagement ($\beta = .31, p < .01$). Meanwhile, a positive attitude towards formal learning was associated with higher levels of affective ($\beta = .18, p < .05$), cognitive ($\beta = .25, p < .05$), and linguistic engagement ($\beta = .22, p < .05$). Together, the attitudes predicted between 9% and 16% of the variation in the five engagement measures (indicating small to medium effects of attitudes). In turn, engagement did not significantly predict any changes in the attitudes towards informal and formal learning over the course of the study. That is, the attitudes at T2 were significantly predicted only by the attitudes at T1.

9.4.2 Motivation, Attitudes, and Engagement Profiles

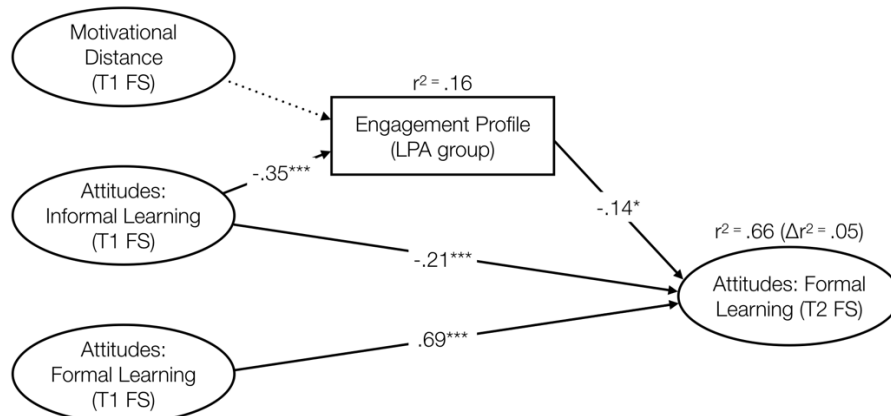
As with proficiency, the researcher also wanted to explore whether motivation and attitudes were linked to participants' general patterns of engagement in ISLPs (*engagement profiles*, as represented by the LPA Groups). First, four separate Factor Score Path Models were run to investigate the relationship of engagement profiles with the (1) language learning orientations, (2) current and future L2 selves, (3) motivational distance, and (4) attitudes towards formal and informal language learning at T1 and T2 (results in Table H.5, Appendix H.2). Subsequently, the variables that emerged from these analyses as sharing significant relationships with the general engagement patterns were entered into one joint model in order to inspect their effects in interaction. However, to avoid issues of multicollinearity between the current and future L2 self and the motivational distance (the latter being directly derived from the former), these variables needed to be entered into separate models with the LPA groups and other motivational variables (Figures 9.24 and 9.25).

Figure 9.24: Current L2 self, attitudes, and engagement profiles, combined Factor Score Path Model. Significant standardised regression path coefficients



Estimation method: Diagonal Weighted Least Squares with pairwise exclusion. $\chi^2_{[3]} = 2.47, p = .480; n = 250; CFI = 1.00, TLI = 1.00, SRMR = .01.$

Figure 9.25: Motivational distance, attitudes, and engagement profiles, combined Factor Score Path Model. Significant standardised regression path coefficients.



Estimation method: Diagonal Weighted Least Squares with pairwise exclusion. $\chi^2_{[3]} = 2.45, p = .484; n = 250; CFI = 1.00, TLI = 1.00, SRMR = .01.$ The dotted line marks a regression path that emerged as significant in the previous analyses but did not reach statistical significance in the current model.

None of the language learning orientations shared any significant relationships with the engagement profiles. In fact, only three variables emerged as significant predictors of LPA group membership. The current L2 self ($\beta = -.19, p < .05$) and attitudes towards informal learning at T1 ($\beta = -.36, p < .01$) were negative predictors of the engagement profiles. This means that participants with a strong current L2 self and positive attitudes towards informal language learning were more likely to be part of LPA Group 1, which showed significantly higher levels of affective

and behavioural and lower levels of linguistic engagement. Motivational distance also emerged as a positive predictor of the engagement profiles when analysed separately ($\beta = .20, p < .01$), but this effect became non-significant in the joint model, suggesting that the same variance was better explained by the attitudes towards informal learning (these variables, in turn, shared a significant negative correlation, $r = -.28, p < .001$).

In the separate models, the informal learning beliefs had the biggest effect on the engagement profiles (partial $r^2 = .17$, medium effect), followed by the current self (partial $r^2 = .09$, small-medium effect) and the motivational distance (partial $r^2 = .04$, small effect). When modelled together, the informal beliefs and current self still explained 20% of the variance in the binary LPA group variable.

The engagement profiles, in turn, significantly predicted changes in only one variable: attitudes towards formal learning ($\beta = -.14, p < .01$). The regression coefficient was negative, indicating that the participants in the second LPA group (who exhibited lower behavioural and affective engagement and higher linguistic engagement) were more likely to 'lose faith' in the effectiveness of formal learning over the course of this study. This effect was small, with the engagement profile accounting for only 5% of the variance not already explained by other variables in the model.

9.4.3 Key Findings

The analyses in the sections above focused on the direct relationships between engagement in informal second language practices and (1) language learning orientations (LLOs, or reasons for wanting to learn English as a second language); (2)

the students' L2 motivational selves (current and ideal future self-perceptions as L2 users); and (3) attitudes towards informal and formal second language learning.

Language learning orientations

Of the five LLOs, only three emerged as significant predictors of engagement in ISLPs: Intrinsic motivation positively predicted (in order of effect size, from stronger to weaker) the levels of linguistic, cognitive and affective engagement. Knowledge orientation positively affected affective and cognitive engagement and the quantity of ISLPs. Finally, instrumental obligation had positive effects on cognitive and affective, but a negative effect on behavioural engagement (both the quantity and variety of ISLPs).

Whereas this was not the case with proficiency (as discussed in the previous chapter), some aspects of engagement in ISLPs did significantly predict changes in some of the LLOs between T1 and T2. In particular, affective and cognitive engagement showed opposite patterns: higher levels of affective engagement and lower levels of cognitive engagement predicted a decrease in intrinsic and a rise in instrumental motivation.

Second language motivational selves

The second set of analyses revealed that the variety of ISLPs was predicted by all three measures of the motivational selves: Students with a stronger current and future L2 self, and a smaller distance between the two, tended to engage in more varied ISLPs. A stronger current L2 self and small motivational distance also predicted more time spent on informal activities (quantity of ISLPs). The amount of variance the variety

and quantity of ISLPs explained by the motivational distance was smaller (4–5%) than what was accounted for by the separate current and future L2 self measures (7–10%).

Different aspects of engagement did, in turn, also predict some changes in the participants' L2 self-concepts between T1 and T2. First, greater variety in ISLPs predicted a small but significant increase in the strength of both the current and future selves at T2 (after controlling for the L2 selves at T1). Meanwhile, a decrease motivational distance between T1 and T2 was predicted by higher levels of affective and linguistic engagement but a lower level of cognitive engagement. However, the three dimensions of engagement could only explain very small proportions of the variance in the motivational selves at T2 (between 1 and 4% beyond what was already explained by the L2 selves at T1).

Attitudes towards informal and formal language learning

Participants' attitudes towards informal learning significantly predicted all dimensions of engagement apart from linguistic engagement. Attitudes towards formal learning were also a significant, but slightly weaker, predictor of affective, cognitive, and linguistic engagement. The effect joint of these two variables was stronger for the quantity of ISLPs ($r^2 = .13$, indicating a medium-sized effect) than the other engagement measures (small effects, r^2 between .6 and .7). None of the individual aspects of engagement significantly predicted any changes in the participants' attitudes between T1 and T2.

Engagement profiles

Finally, the relationships between the three sets of motivational and attitudinal variables and the engagement profiles of the LPA groups were also investigated.

These analyses revealed that the students in Group 1 (who showed higher levels of behavioural and affective and lower levels of linguistic engagement) had significantly more positive perceptions of their own current L2 competence and more positive attitudes towards informal learning. There was also a trend in Group 2 (who engaged less in ISLPs but showed higher levels of linguistic engagement) to grow slightly more negative in their attitudes towards formal language learning over the course of this study.

9.5 Research Question 4: How do General Language Proficiency, Language Learning Motivation, and Attitudes Interact in their Relationships with Engagement in Informal Second Language Practices?

The models in Sections 9.3 and 9.4 focused on the direct relationships between engagement in informal second language practices and general language proficiency, motivation for language learning, and attitudes towards informal and formal learning. Sections 9.3 and 9.4 were dedicated to exploring the direct relationships between engagement in informal second language practices and general language proficiency, motivation for language learning, and attitudes towards informal and formal learning. The analyses described in Section 9.5 build on these findings. By formulating models which combine those variables that were found to share significant relationships with any aspect of engagement into the same models, the analyses presented here explore the interactions between proficiency, motivation, and attitudes in relationship to engagement in informal second language practices. Separate models were formulated to investigate the relationships between these variables and (1) the five measures of engagement in ISLPs (Section 9.5.1) and (2) the LPA groups, coding for general engagement patterns (Section 9.5.2).

9.5.1 Behavioural, Affective, Cognitive, and Linguistic Engagement

TO examine the relationships between proficiency, motivation, attitudes, and the different dimensions of engagement, three separate models were formulated. This was because entering all variables into the same model would have made this analysis too complex. In addition, motivational distance had to be modelled separately from the current and future L2 selves (as previously in Section 9.4.1.2) to avoid multicollinearity issues. Thus, alongside proficiency and the five engagement variables, the LLOs were included Model 1; the current and future L2 selves and attitudes towards informal and formal language learning in Model 2; and the motivational distance and attitudes in Model 3.

As mentioned previously (Section 2.3), there is disagreement about the directionality of the relationship between motivation and proficiency: It may be that more motivated students tend to become more proficient or that more proficient students are more motivated to further their language development. Indeed, these two variables might be also be mutually reinforcing. In the current study, however, the researcher chose to enter proficiency into the Path Models before the motivational variables, both at T1 and T2. This was because the proficiency test was administered before the motivation and attitudes questionnaire, and because it seemed plausible that a participant's actual proficiency may colour the ways in which they perceived their L2 and themselves as a language learner at any given time. Across the two time points, however, both the impact of proficiency on changes in motivation and vice versa were modelled.

Altogether, then, the tested models included five *groups of variables*, each of which were regressed on the previous ones: proficiency at T1, motivation and

attitudes at T1, engagement in ISLPs, proficiency at T2, and motivation and attitudes at T2. Only paths to and from the engagement variables were included in the models, which had emerged as significant in the previous analyses (Sections 9.3 and 9.4). Constructs which were not previously found to share significant links with engagement were not included in the model (e.g. international orientation and instrumental benefits at both T1 and T2; knowledge orientation at T2). The significant results of the three Factor Score Path Models are presented in Figures 9.26 to 9.28 (p. 253–255) and reported in full in Tables I.1–I.3, Appendix I.1.

Including both proficiency and the motivational and attitudinal variables in the same model with engagement resulted in a number of changes to the magnitude, and in some cases the significance, of the regression paths. The changes to the relationships of proficiency, language learning orientations, the L2 motivational selves, and attitudes towards language learning are discussed in turn below.

Proficiency

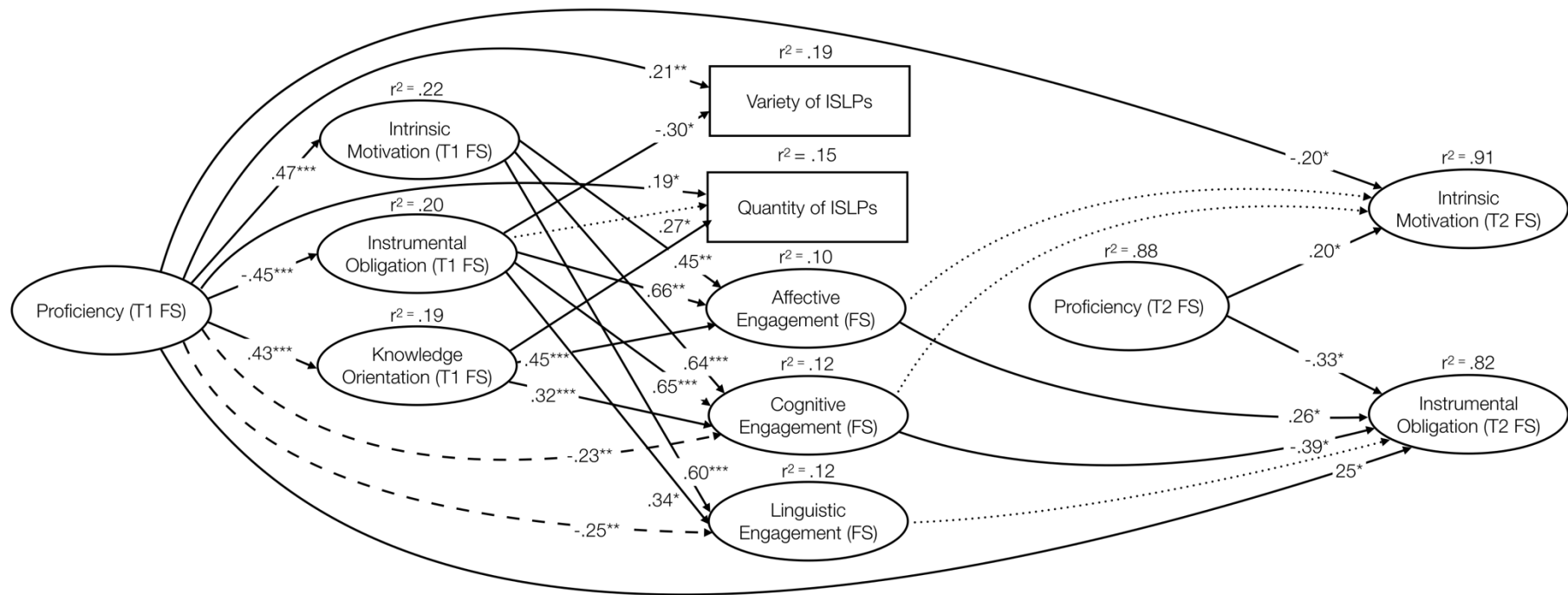
The regression coefficients representing the direct effects of proficiency on the variety and quantity of ISLPs were smaller in all three of the new models than in the model which did not include motivation and attitudes (Section 9.3). This suggests that motivation and attitudes, to a certain extent, mediated the relationship between proficiency and engagement. This mediation effect was strongest in the second model, which included the current and future L2 selves and attitudes towards language learning (Figure 9.27), followed by the language learning orientations model (Figure 9.26) and the model including the motivational distance and language learning beliefs (Figure 9.28).

As mentioned above, only those paths were included in these models which emerged as significant in the earlier stages. However, the modification indices were also consulted at this point to check whether any significant relationships were being overlooked in the model. Indeed, the MI indicated a direct effect of proficiency at T1 on cognitive engagement (Figures 9.26–9.28, dashed lines), which had not been not significant in the previous model without the motivation constructs (Section 9.3). Additionally, the path from prior proficiency to linguistic engagement, which had only been marginally significant in the previous analysis, was now significant at $\alpha = .05$.

The fact that the effects of proficiency on cognitive and linguistic engagement emerged as significant only when adding language learning motivation and attitudes to the same model can be explained by looking at the directionality of these effects: Proficiency at T1 *negatively* affected both cognitive and linguistic engagement. That is, more proficient participants tended to focus less on the informal activities themselves, as well as on the specific language features contained within them (direct negative effects).

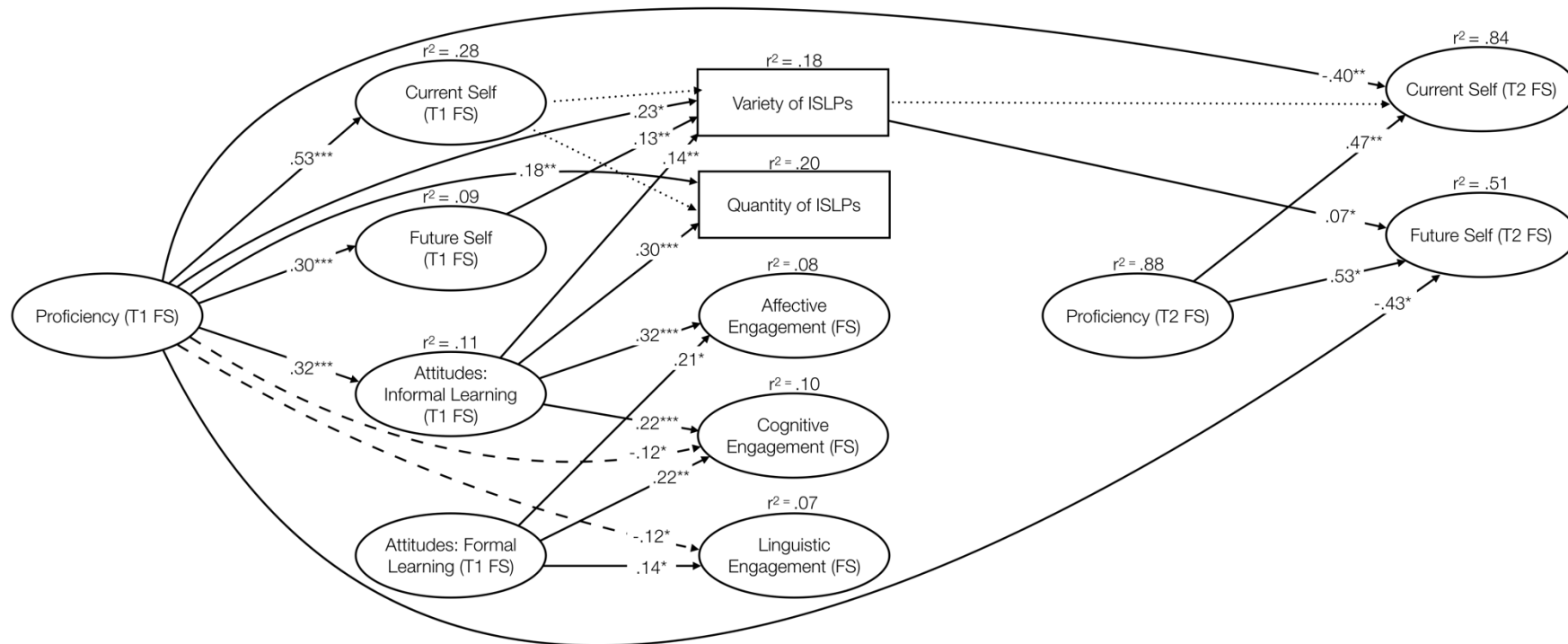
Meanwhile, there were also several indirect positive effects of proficiency on cognitive and linguistic engagement via some of the motivation and attitudes variables. In particular, more proficient participants tended to have a stronger intrinsic motivation, knowledge orientation, and more positive attitudes towards informal learning, which in turn predicted higher levels of focus on the ISLPs activities themselves (cognitive engagement) and the language therein (linguistic engagement).

Figure 9.26: Proficiency, language learning orientations, and engagement, Factor Score Path Model. Significant regression path coefficients



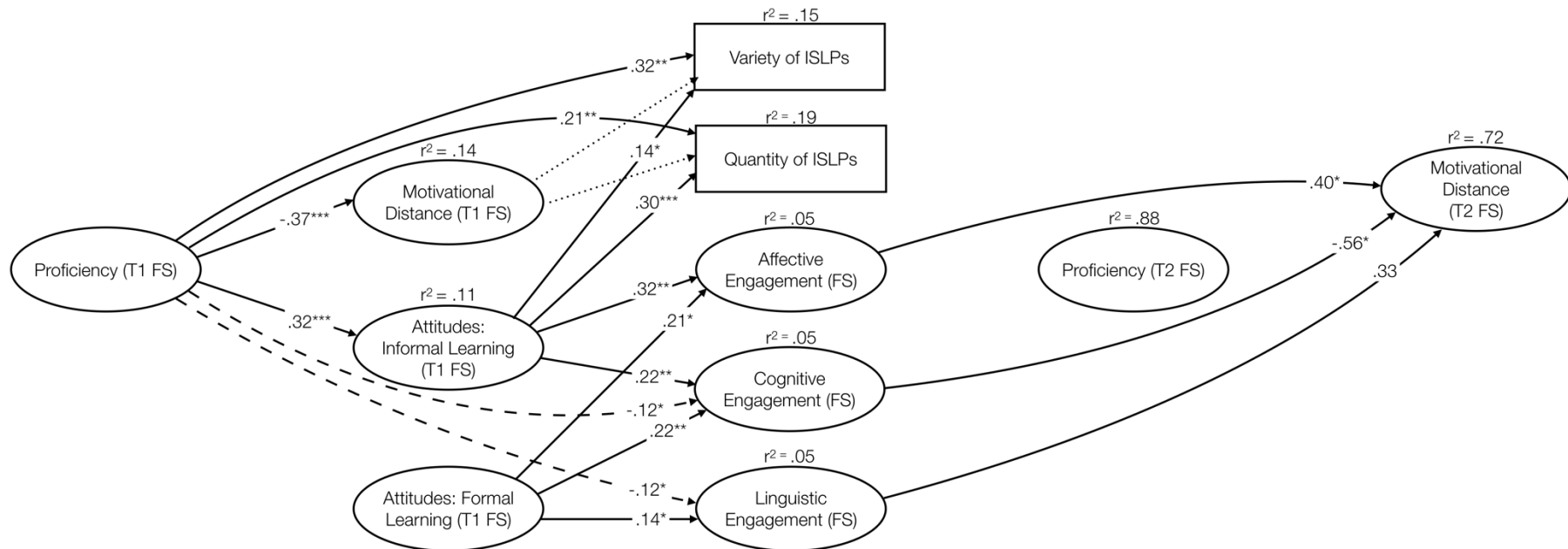
Note: Not pictured are the significant paths from proficiency and motivation at T1 to T2. Dotted lines mark regression paths that emerged as significant in the previous analyses but did not reach statistical significance in the current model. Conversely, dashed lines represent relationships which were non-significant in the previous models but reached significance in this analysis. Estimation method: Full Information Robust Maximum Likelihood. $\chi^2_{[9]} = 15.58, p = .076; n = 331; CFI = .99, TLI = .98, SRMR = .02.$

Figure 9.27: Proficiency, motivational selves, attitudes towards language learning, and engagement, Factor Score Path Model. Significant regression path coefficients



Note: Not pictured are the significant paths from proficiency, motivation and attitudes at T1 to T2. Dotted lines mark regression paths that emerged as significant in the previous analyses but did not reach statistical significance in the current model. Conversely, dashed lines represent relationships which were non-significant in the previous models but reached significance in this analysis. Estimation method: Full Information Robust Maximum Likelihood. $\chi^2_{[28]} = 36.59, p = .128; n = 331; CFI = .99, TLI = .98, SRMR = .04$.

Figure 9.28: Proficiency, motivational distance, attitudes towards language learning, and engagement, Factor Score Path Model. Significant regression path coefficients



Note: Not pictured are the significant paths from proficiency, motivation, and attitudes at T1 to T2. Dotted lines mark regression paths that emerged as significant in the previous analyses but did not reach statistical significance in the current model. Conversely, dashed lines represent relationships which were non-significant in the previous models but reached significance in this analysis. Estimation method: Full Information Robust Maximum Likelihood. $\chi^2_{[28]} = 36.59, p = .128; n = 331; CFI = .99, TLI = .98, SRMR = .04$

In the previous model, which did not include the motivational variables (Section 9.3), the direct negative and indirect positive effects of proficiency on cognitive and linguistic engagement would have been attributed to the same path and thus seemed to have out cancelled each other (resulting in non-significant regression paths). This illustrates the complexity of the relationships between these variables and the importance of separately modelling direct and indirect effects.

The regression coefficients of the direct paths from proficiency to cognitive and linguistic engagement were higher in the first model ($\beta = -.22$ and $-.25$ respectively) than in the second and third (all $\beta = -.12$), indicating that these relationships were more strongly mediated by the motivational selves and learning beliefs than by language learning orientations (just like the effects of proficiency on the two behavioural engagement measures).

Language learning orientations

Concurrently modelling the effects of proficiency and motivation also resulted in several changes to the direct relationships between the language learning orientations and engagement variables. First, there were changes to the strength of several of the effects of LLOs on engagement: When proficiency was modelled alongside these variables, the direct effects of intrinsic motivation and knowledge orientation on several of the engagement dimensions gained in strength (larger β -values). Meanwhile, the direct effects of instrumental obligation were less pronounced (smaller β -values). This suggests that intrinsic motivation was a better predictor of engagement than proficiency, whereas some of the variance previously attributed to instrumental obligation was better explained by proficiency at T1.

Furthermore, a number of previously significant regression paths (Section 9.4.1.1) became non-significant after the inclusion of proficiency in the same model (Figure 9.26, dotted lines). On the one hand, the effect of instrumental obligation on the quantity of ISLPs was no longer significant, indicating that the same variance was better explained as a direct effect of proficiency. On the other hand, the paths from affective and cognitive engagement to intrinsic motivation at T2, as well as from linguistic engagement to instrumental obligation at T2 became non-significant, indicating that these effects may have emerged in the previous model due to an indirect effect of proficiency on changes in these LLOs.

Second language motivational selves

Both of the previously significant effects of the current L2 self on the quantity and variety of ISLPs (Section 9.4.1.2) no longer reached significance in the combined model (Figure 9.27, dotted lines). The same was true for the effects of the motivational distance on the two measures of behavioural engagement (Figure 9.28, dotted lines). With proficiency being a strong predictor of both the current L2 self (which arguably measures, in part, learners' self-estimated proficiency and self-efficacy) and the motivational distance, it seems that these constructs had been accounting for the same variance in behavioural engagement in the previous, separate direct effects models. The ideal future self was the only part of the motivational L2 self-system which was more or less unaffected by the addition proficiency to the same model.

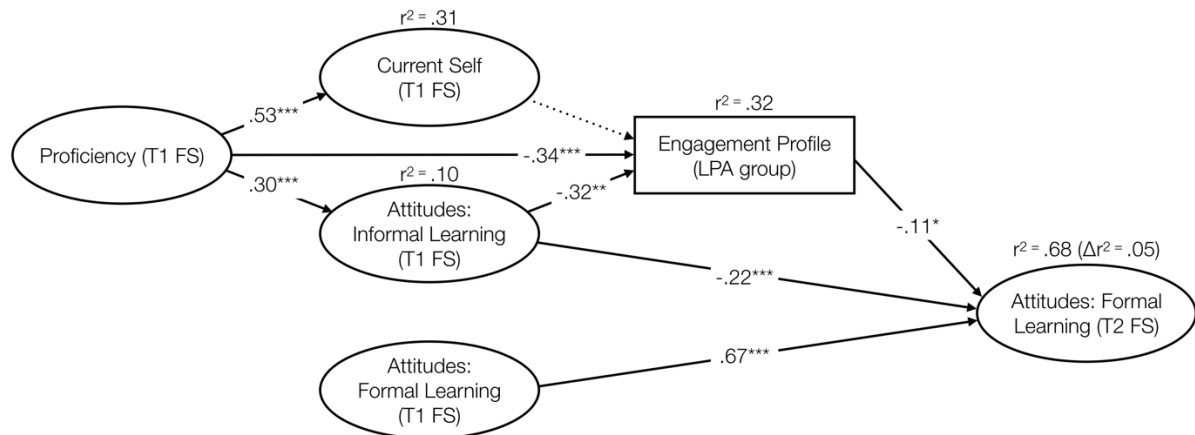
Attitudes towards informal and formal learning

Finally, the direct relationships between the attitudes towards informal and formal language learning and engagement were not significantly different when modelled alongside the effects of proficiency and the motivational L2 selves. The regression coefficients representing the effects of attitudes towards informal learning on engagement were slightly smaller in these models (Figures 9.27 and 9.28) than before (Section 9.4.1.3). This indicates that some of the variance in engagement previously attributed to attitudes towards informal learning was already accounted for by proficiency at T1, which in turn also had a significant effect on attitudes towards informal learning beliefs, indicating the presence of a mediation effect. By contrast, the relationships between the attitudes towards formal learning and engagement were essentially unaffected, presumably because these attitudes themselves were not significantly predicted by language proficiency at T1.

9.5.2 Engagement Profiles

Finally, a Factor Score Path Model was formulated which included proficiency alongside motivation, attitudes and the engagement profiles extracted (LPA groups). As in the analyses presented in the previous section, only paths which had emerged as significant in the prior modelling steps (Sections 9.3.4 and 9.4.2) were included in this model and proficiency was entered ahead of the motivation and engagement variables. The results are reported in Table I.4, Appendix I.2 and in Figure 9.29.

Figure 9.29: Proficiency, motivation, attitudes, and engagement profiles, combined Factor Score Path Model. Significant standardised regression path coefficients



Estimation method: Diagonal Weighted Least Squares with pairwise exclusion. The dotted line marks a regression path that emerged as significant in the previous analyses but did not reach statistical significance in the current model. $\chi^2_{[2]} = 3.19, p = .074; n = 235; CFI = .99, TLI = .93, SRMR = .02$.

Concomitantly modelling the relationships of proficiency, the current L2 self and language learning attitudes with the engagement profiles resulted in only one significant change: The effect of the current self at T1 on LPA group membership was no longer significant (Figure 9.29, dotted line). Again, this shows that a proportion of the variance in the engagement profiles previously attributed to the current L2 self was better explained by proficiency. Besides this, the simultaneous modelling changed the magnitude of the effect of proficiency on the engagement profile (previously $\beta = -.47$, now $\beta = -.34$). This means that some of the variance in binary LPA group variable previously attributed directly to proficiency at T1 was better explained as an indirect effect via the attitudes towards informal learning. The simultaneous modelling did not substantially affect the magnitude of the other regression coefficients.

Altogether, proficiency and attitudes towards informal learning had a large effect on the engagement profiles, explaining 32% of the variance. In contrast, the

effect of the general engagement patterns on the attitudes towards formal learning at T2 was small (partial $r^2 = .05$).

9.5.3 Key Findings

The three models discussed in Section 9.5.1 addressed the interplay between proficiency, motivation, and attitudes in relation to engagement in ISLPs. Respectively, these models included proficiency and the five measures of different engagement dimensions alongside (1) the language learning orientations; (2) the current and future L2 selves and attitudes towards informal and formal language learning; and (3) the motivational distance and attitudes.

In all three models, the direct effects of proficiency on engagement were weakened by the inclusion of the motivation and attitudes measures, suggesting that these variables, separately, had explained some of the same variance in engagement. The mediation effects that were detected in this way were strongest in Model 2, followed by Model 1 and Model 3. Furthermore, the modification indices showed that these models (which at first included only the paths which had emerged as significant in the analyses presented in Sections 9.3 and 9.4) were missing some significant relationships between the specified variables: In particular, it emerged that proficiency at T1 had significant negative effects on cognitive and linguistic engagement (the latter effect had only been marginally non-significant before, see Section 9.3.3). The fact that these direct effects of proficiency on engagement only now emerged as significant suggests that they were cancelled out by indirect effects in the opposite direction (via motivation) in the previous models. This illustrates the complexity of the relationships between engagement in ISLPs, proficiency,

motivation, and attitudes and emphasises the importance of modelling these constructs together, rather than separately.

Looking at the motivation variables, the most noticeable changes caused by the inclusion of proficiency in the same models were with regard to the current L2 self. This variable was no longer a significant predictor of behavioural engagement (quantity and variety of ISLPs). This might be attributed to the close relationship between proficiency and the current L2 self (which could be viewed as a partial measure of self-estimated proficiency or self-efficacy). Presumably for the same reasons, the effects of motivational distance at T1 (which was a measure directly derived from the current L2 self) on behavioural engagement also became non-significant after the inclusion of proficiency in the same model. The relationships of the future self with engagement, however, were more or less unaffected.

Concurrent modelling with proficiency affected only a few of the direct paths between the language learning orientations and engagement: First, instrumental orientation (which itself was strongly negatively affected by proficiency at T1) was no longer a significant predictor of the quantity of ISLPs. Second, the effects of both affective and cognitive engagement on intrinsic motivation at T2 became non-significant, indicating that these changes in motivation were better explained by the concomitant changes in proficiency from T1 to T2 rather than engagement in ISLPs.

Finally, the inclusion of proficiency in the same model did not much affect the relationships between the attitudinal variables and engagement: The effects of attitudes towards informal and formal learning on engagement were slightly weaker in these models, but all of them remained statistically significant.

The interplay of the motivational and attitudinal variables and proficiency was also modelled in relation to the engagement profiles (LPA groups). Similar changes appeared in this model as those described above: The direct effect of the current L2 self at T1 on the general engagement patterns no longer reached significance after the inclusion of proficiency at T1. What remained were significant effects of proficiency and attitudes towards informal learning: More proficient students, who also tended to believe more in the effectiveness of informal learning, were more likely to be part of the first LPA group, which exhibited higher levels of behavioural and affective, but lower levels of linguistic engagement. Meanwhile, after controlling for the effects of proficiency, the engagement profiles still predicted a slight change in the attitudes towards formal learning.

10. Discussion

In this chapter, the findings of the current study will be discussed, focusing on each of the following questions in turn:

1. How do learners of English as a second language, attending year 10 classes at German *Gymnasien*, engage in English informal second language practices? (Section 10.1)
2. What are the relationships between general language proficiency and engagement in informal second language practices? (Section 10.2)
3. What are the relationships between motivation and attitudes towards language learning and engagement in informal second language practices? (Section 10.3)

Wherever possible, the results of the quantitative data (Phase III) will be related to insights from the qualitative analysis (Phase I). In addition, the findings are also discussed in relation to the outcomes of previous research on informal second language learning.

10.1 Engagement in Informal Second Language Practices

The current study was the first to investigate the informal second language practices of secondary school students in Germany (15–16 years old, attending year-10 classes at *Gymnasien*). Therefore, one aim was to discover which types of informal activities were most popular among these students and how much time they spent on them (see Section 10.1.1). However, in addition to these aspects of behavioural engagement, this study also investigated the learners' thoughts and feelings while participating in these informal second language practices; that is, the extent to which they were affectively, cognitively, and linguistically engaged (Section 10.1.2).

10.1.1 Behavioural Engagement

In addition to any instruction in school (2.5 hours/week) and homework (c. 2 hours/week), the students in this sample engaged in English-language activities informally for an average of 108 min per day (c. 1.8 hours/day, or 12.6 hours/week). There was vast individual variation in the time dedicated to ISLPs, although just over half of the scores fell between 30 and 120 minutes (0.5 to 2 hours/day). In general, individual participants' informal practices were not particularly varied, with most of them engaging in only one or two unique activities on any given day. This does not mean that they engaged in the same activities every day, but that there was little variation within one day.

There was a strong, positive correlation between the quantity and variety of students' informal activities. That is, those who spent more time on ISLPs per day also tended to have more varied repertoires of informal practices. Looking at individual diary entries in more detail, the researcher was also able to deduce that many of the participants who recorded the highest daily duration of ISLPs spent long, continuous amounts of time immersed in only one activity, most often listening to music, watching serials or YouTube videos, or playing video games. Meanwhile, the pupils who spent the least time on ISLPs per day (less than 20 min) tended to engage very briefly either in listening to music (i.e. only a few songs) or with social media (e.g. checking their timelines). These findings can be triangulated with insights from the qualitative data: In the focus groups, many of the participants who gave the highest estimates of how much time they typically spent on ISLPs (3–4 hours/day or more) told the researcher that they frequently listened to music, played video games, or 'binge-watched' serials or online videos for several hours. By contrast, participants who estimated that they spent very little time on ISLPs (less than 1 hour/day) often

claimed that their informal contact was limited to English songs they encountered by chance, for example on the radio, or occasionally looking at social media posts.

Across the whole sample, the most popular activities were listening to music, using social media (including watching videos on YouTube), and watching serials. An important finding was the sharp difference in the prevalence of these activities: While about 80% of the participants reported listening to music or using social media at least once during the study, less than one third of them watched serials English. This is far lower than the proportion of students in other studies who reported regularly watching English-language TV series (e.g. Kuppens, 2007, 2010; Sockett, 2014d; Toffoli & Sockett, 2010, see Section 2.1.1). The least popular practices were listening to spoken English (e.g. audiobooks, radio programs, conversations taking place in the vicinity), watching films, and writing, which included short text message exchanges, and producing longer texts (e.g. emails, blog posts, creative writing in English). Less than 15% of the sampled participants engaged in any of these activities.

Furthermore, three findings bear mentioning here, which stood out when examining how exactly language was used in the informal practices: First, the number of entries pertaining to 'listening to music while singing along' was almost the same as those reporting only listening (Appendix F.1.5.). This does not necessarily mean that the pupils were singing along all the time, but it suggests that, for these students, listening to music often involved some language *reproduction* and was thus not a purely receptive activity. The same is true of playing video games: In approximately 50% of the entries pertaining to multi-player gaming, participants also reported communicating with other players through text and/or voice chat (Appendix F.1.6). Finally, it was revealed that watching videos online (primarily on YouTube) was

actually more popular among the sampled students than watching serials (including through online streaming services such as Netflix). This finding lends some support to the claim, frequently discussed in online news media (e.g. Parrish, 2016; C. Williams, 2016), that YouTube is increasingly ‘replacing television’, especially, but not exclusively, among the younger generations.

10.1.1.1 Gender Differences

Even though gender differences in ISLPs were not an explicit focus in this study, some interesting findings emerged with regard to the types of informal practices in which boys and girls predominantly engaged. In the focus group interviews, the researcher observed that a majority of the girls, but none of the boys, mentioned reading short texts and watching videos in English on Instagram. Meanwhile, all interviewed boys played video games, and 60% concurrently chatted with other players in English, whereas only 20% of the girls played video games and only one communicated with other players. Triangulating these observations with the quantitative data revealed that boys were indeed more likely than girls to play video games and even more so to communicate via text or voice chat with fellow players. Girls, on the other hand, were more likely to engage in all of the other informal activities except for purely video- and text-based social media use, where there was no significant gender difference. Yet, there was only one significant difference in the *time* that boys and girls spent on different activities: boys spent significantly more time using video-based social media, but not text-based and mixed social media. Follow-up analyses revealed that these findings could be explained via differences in the particular social media platforms that were popular among boys and girls: In particular, girls were more likely than boys to use Instagram, and thus to engage in mixed social media consumption,

whereas boys spent significantly more time on YouTube, and thus on video-based social media.

A number of prior studies found significant gender differences in ISLPs, which are sometimes also associated with variations in proficiency (Kuppens, 2007, 2010; Sundqvist, 2009, 2011; Sundqvist & Sylvén, 2014). However, in this study, a comparative analysis indicated that the boys and girls in this sample did not differ significantly with regard to their overall levels of behavioural engagement (i.e. the total time spent per day on informal practices and the variety of activities) and affective, cognitive and linguistic engagement, nor regarding their proficiency at either point of data collection. Nevertheless, it would have been interesting to also explore whether gender affected the relationships between ISLPs, proficiency, motivation, and attitudes that were the focus of this study. However, conducting separate analyses by gender for each of the subsequent research questions (Sections 9.3–9.5) would have gone beyond what was possible in the scope of this thesis, in terms of the size of the data set required for such analyses and the available project time. Thus, whether differences in the engagement in ISLPs could explain some of the gender differences uncovered in previous studies with other learner populations remains a topic for future research.

All in all, the findings regarding the behavioural engagement of the students in this sample (the quantity and type of informal practices in which they participated) roughly replicated what has been found in prior research: For example, most previous studies also reported that most of their participants came into contact with English in their free time for close to two hours per day (e.g. Cole, 2015; Lai et al., 2015; J. S. Lee, 2019; Olsson, 2011). Just as in the current study, the majority of this time was

generally dedicated to receptive rather than productive activities, and listening to music, watching serials and online videos were among the most popular practices (e.g. Lamb & Arisandy, 2019; Lyrigkou, 2018; Sockett, 2014d). Other researchers have also observed that video games were significantly more popular among boys (e.g. Kuppens, 2010; Olsson, 2011; Sundqvist & Sylvén, 2014), although the finding of some studies (Olsson, 2011; Sundqvist & Sylvén, 2014) that boys have more informal contact with English overall was not replicated here.

Nevertheless, only rough comparisons can be made between the current study and prior research which described the informal second language practices of other learner populations. This is because any differences in findings could be attributed to a variety of factors, including effects of the age and experience of the learners as well as the cultural and educational context of the studies, but also differences in the data collection methods used (e.g. general and retrospective self-report questionnaires or language learning diaries, sometimes measuring the frequency of participation in different activities and sometimes the time spent on them). Thus, to be able to draw valid conclusions about group differences and the effects of the cultural and learning context, larger cross-sectional and cross-national studies, employing unified measures of engagement in informal second language practices, will need to be conducted in future. In addition, the field could also benefit from more longitudinal studies which track developments in ISLPs in the same population over time (cf. Olsson & Sylvén, 2015).

10.1.2 Affective, Cognitive, and Linguistic Engagement

Several researchers have proposed that the effort which L2 learners put into certain informal activities (i.e. how cognitively and/or linguistically engaged they are) might contribute to differences in the language development associated with ISLPs (e.g. Cole, 2015; Sundqvist & Sylvén, 2016). However, with one exception (Lyriqkou, 2016, 2018, see Section 2.2.2), there have been no studies which have included any measures of learner effort in informal activities. Therefore, in Phase III of the current study, the researcher asked the participants to self-report their levels of affective, cognitive, and linguistic engagement after each informal practice. The findings from these data are discussed below, contextualised with insights from the focus group interviews. On the whole, they show that the extent to which learners were mentally and emotionally engaged did vary, to some extent, according to the type of activity in which they are participating.

Affective engagement

The qualitative data revealed that affective engagement was one of the main driving forces behind participation in ISLPs in the sampled population: It seemed that informal English-language activities had to evoke a certain level of interest and potential enjoyment in the participants to be worth doing (in addition to not being too challenging, see *cognitive engagement* below). This is also reflected in the quantitative data: The distributions of all three items which measured affective engagement in the informal language learning diary were negatively skewed, meaning that, in most activities, learners reported high levels of interest and enjoyment. This also helps to explain why the average levels of affective engagement did not substantially between most of the different types of informal activities.

Nevertheless, the activity of *playing video games and chatting with other players* was associated with significantly higher affective engagement than all other informal practices, followed, at some distance, by watching serials and playing video games *without* concurrent communication. The high emotional investment in video games, which seemed only to be strengthened through social interactions with other players, is not altogether surprising: The immersive nature of video games has been discussed extensively in the literature and has led to suggestions of using games as engaging tools for learning and teaching (e.g. de Freitas, 2013). The high levels of affective engagement evoked by serials be emblematic of pupils becoming, over time, increasingly immersed in the programmes' narratives and emotionally invested in the lives of the main characters.

Cognitive engagement

In addition to affective engagement, the qualitative data also suggested that the required levels of cognitive engagement determined, to some extent, whether or not individuals chose to take part in certain informal activities. That is, most of the focus group participants who reported engaging in ISLPs often said that they were able to understand the language in these activities without much effort; some even claimed that they perceived no difference between watching a video, reading a text, etc. in English and German. By contrast, those participants who did not frequently engage in ISLPs said that this was in part because they had to concentrate too much on each activity to reach a level of comprehension that would allow them to enjoy it. A similar finding was previously reported in a study of Hong Kong and US-American L2 learners' out-of-class practices by Lai, Hu, and Lyu (2018, see Section 2.3.2), where the frequency of engagement in entertainment- and information-oriented practices

(roughly equivalent to what is covered by the term *informal second language practices* in the current study) was significantly predicted by the perceived ease of such activities.

There was some disagreement among the participants over which types of informal activities were the most challenging: In particular, a number of pupils described struggling with watching serials or online videos because these activities required them to quickly process large amounts of spoken language, leaving no time to focus on individual unknown phrases without losing the overall narrative thread. Meanwhile, many others commented on the exceptional focus that reading books required, in part due to the narrative complexity, but also the fact that, when one is reading, comprehending the meaning depends solely on understanding the language (i.e. one cannot make inferences from the visual context, such as when watching videos or looking at pictures on social media). Indeed, reading books was among the activities which, according to the quantitative data, was associated with the highest average levels of cognitive engagement, together with the only three types of activities that involved language production: writing in English, speaking, and communicating with fellow players while playing video games. The latter finding was expected, given that language production, in contrast to receptive skills, is generally regarded as more difficult by learners, teachers, and researchers alike. More importantly, however, this finding corroborates Sylvén and Sundqvist's (2016) proposed distinction between 'active' and 'passive' informal practices (see Section 2.3.2): Namely, the same activities which emerged as requiring greater cognitive engagement in this study — speaking, writing, reading, and playing video games — were labelled as 'active' by Sylvén and Sundqvist, who judged that these practices required learners to be more active and rely more strongly on their language skills.

Linguistic engagement

While the findings presented above suggest that cognitive engagement is, in part, related to language processing, this study also explored separately the dimension of *linguistic engagement*, that is the ‘attention to form’ in informal activities. A vast majority of the focus group participants reported that they did not generally focus much on the language in ISLPs. This was also reflected in the quantitative measures, as the distributions of the three linguistic engagement items had mean scores towards the lower end of the scale, well below those for cognitive and affective engagement.

Nevertheless, some interviewed pupils did recount specific circumstances under which they might choose to focus on a particular linguistic feature. For example, if a word or phrase they did not know was integral to the plot of a series or book or repeated often within it, they might dedicate some effort towards understanding it. However, almost all pupils stressed that their primary purpose in doing so would be to increase their understanding of the overall *meaning* and not to improve their language skills.

Altogether, these findings support the conclusion that these students — much like the participants in prior studies (e.g. Olsson, 2011; Sockett, 2014b; Sundqvist, 2009, 2011; Sundqvist & Olin-Schneller, 2013) — engaged in ISLPs mostly for entertainment purposes (cf. *affective engagement* above) and were not strongly motivated by a desire to improve their language proficiency. Other studies, too, have concluded that most ESL learners do not intentionally and/or strategically engage with unknown language features in informal second language activities (Cole, 2015; Lyrigkou, 2016). This suggests that any language learning which is found to be associated with informal second language practices is likely to be mostly incidental.

Although the general levels of linguistic engagement were rather low, learners were significantly more likely to pay conscious attention to the linguistic features in some informal practices than others. Specifically, writing, speaking, and reading were associated with the highest levels of linguistic engagement. This is probably because these activities are entirely focused on language processing, unlike multimodal practices such as watching serials or online videos or browsing social media, which involve both a visual and a language component.

Since the learners also reported some of the highest overall levels of focus during reading, writing and speaking, the cognitive and linguistic engagement dimensions seem to be closely interrelated. However, one finding in particular illustrated that these two factors were, indeed, conceptually distinct: Whereas *playing video games while communicating with other players* was among the activities with the highest levels of cognitive engagement, the associated levels of linguistic engagement were significantly below all other activities. This suggests that, while playing games and connecting with other players, the participants were highly focused (cognitively engaged) not because this activity involved language production (as postulated above), but because it required the parallel coordination of many tasks (in addition to producing language), including active gameplay, strategic planning, and in some cases cooperation with other players.

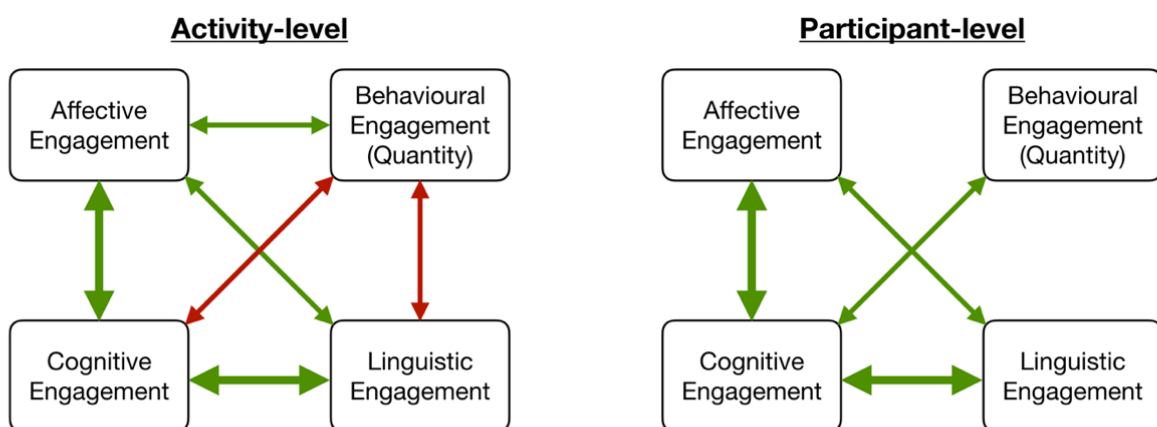
Thus, this particular finding is valuable in supporting the theoretical distinction between cognitive and linguistic engagement as separate, but closely interrelated dimensions (which also emerged in the factor analyses, Sections 7.3 and 9.1.3.5): It seems that cognitive engagement captured the overall strength of mental effort that learners dedicated to an activity, while linguistic engagement measured how much of that effort was directed towards consciously decoding linguistic forms. Thus,

linguistic engagement appears to be predicated on cognitive engagement: If a learner is not attentive while engaging in an activity (such as when listening to music only in the background, while also doing their homework), it is less likely that any particular language feature would become salient enough to command their focused attention. Yet, as the case of *playing video games while chatting with other players* shows, high levels of cognitive engagement do not always result in conscious attention to form.

10.1.3 Relationships between Engagement Dimensions

Following on from the previous paragraph, at this point the discussion will turn towards the relationships between the different dimensions of engagement measured in this study. Several such links emerged in the qualitative data, which will be compared to the correlations between the different factors in the quantitative data (depicted in Figure 10.1). By using Multi-level Confirmatory Factor Analysis (MCFA; Section 9.1.3.5) to analyse these data, the researcher was able to separately investigate the correlations between these variables at the *activity-level* (within the same reported activity) and the *person-level* (individual tendencies in engagement).

Figure 10.1: Relationships between the four engagement dimensions at the activity- and participant-level.



Note: Green arrows indicate positive and red arrows show negative relationships. The width of the arrows denotes the strength of the relationships (Pearson's $r < .30$ vs. $r < .50$).

As mentioned before, the qualitative data indicated that enjoyment and interest were the main determinants of whether or not participants chose to engage in informal second language practices. Indeed, the activity-level data also showed that participants spent slightly more time on activities from which they derived greater enjoyment. However, affective and behavioural engagement did not significantly correlate at the participant-level, that is, participants who spent more time on ISLPs did not necessarily experience greater levels of enjoyment and interest. This implies that affective engagement influences the decision to engage in informal practices situationally. In other words, learners seem to decide to engage in ISLPs based on their interest in a specific activity and the enjoyment they might derive from it, rather than according to whether they *generally* enjoy using their L2 informally.

The qualitative data also suggested that affective engagement might, to some extent, determine the pupils' concurrent cognitive and linguistic engagement. That is, the interviewed students reported that they were more likely to focus closely on an informal activity (cognitive engagement) and to pay conscious attention to the language within (linguistic engagement) if they were enjoying themselves and were truly interested in the activities (affective engagement). In the MCFA, affective engagement was indeed strongly correlated with cognitive engagement at both the activity- and participant level. However, the direct correlations between affective and linguistic engagement were only small, but still significant.

A possible explanation for these latter, smaller correlations is that the suspected effect of affective engagement on linguistic engagement might have been mostly *indirect*, via cognitive engagement. This is supported by the qualitative finding that cognitive engagement seemed to be a precursor to linguistic engagement (Section 9.2.4), and the fact that these two factors were highly correlated at both

levels of the MCFA. Another possibility is that the expected positive effect of affective engagement on linguistic engagement was weakened by a concurrent negative effect of linguistic engagement on affective engagement, resulting in an overall smaller correlation coefficient: While most of the focus group participants agreed that they would only be willing to try and decode unknown words, phrases, or constructions if they were really interested in an activity (positive effect of affective engagement on linguistic engagement), they also noted that *having to* focus on trying to understand particular language features would negatively affect their enjoyment of any activity in which they engaged (negative effect of linguistic engagement on affective engagement). Presumably, the latter effect would have predominantly been reported by students with lower English proficiency, who would be more likely to encounter more comprehension issues. However, this theory cannot be verified directly, as no data about the focus group participants' proficiency was collected in Phase I (however, see Section 10.1.4).

Another set of relationships that needs further explanation are those of cognitive and linguistic engagement with behavioural engagement. In the MCFA, the duration of an activity was found to be shared weak negative correlations with both cognitive and linguistic engagement, suggesting either that learners were losing their focus (on both the activity itself and the language) over time, or that they spent less time on activities which required more concentration. By contrast, at the participant-level, there was a moderate *positive* correlation between cognitive engagement and the quantity of informal activities, which indicated that participants who generally spent more time on ISLPs also tended to pay closer attention to them. In trying to explain this contrast, the researcher examined whether specific activity types, which required higher levels

of focus, were particularly popular among participants who spent a lot of time on ISLPs in general. Indeed, of the 29 participants (10%) who recorded the highest quantity of engagement in informal activities (more than 1.5 hours/day), ten *played video games and chatted with other players* for more than 50 minutes per day, which was the activity with the highest average level of cognitive engagement. No such patterns were found, however, regarding reading, writing, and speaking in English, the other activities associated with higher cognitive engagement scores.

Meanwhile, linguistic engagement was not significantly correlated with the ISLL quantity at the participant-level, although there was a small but significant negative correlation at the activity-level, as mentioned above. As with affective engagement (discussed earlier in this section), the differences between the correlations at the activity- and participant-level could be interpreted as showing that the degree of linguistic engagement influenced the decision of these learners to engage in informal practices *situationally*: That is, the learners might have been discouraged from spending more time on a particular activity if, in order to comprehend it, they needed to pay very close attention to the language. Nevertheless, the lack of a correlation at the participant-level means that the pupils did not *generally* seem to shy away from ISLPs that required some linguistic focus, or, conversely, that those students who engaged more in ISLPs did not generally focus less on form.

This result is in direct contrast to a previous finding by Lyrigkou (2016), who reported a significant *negative* correlation between the ‘mental effort’ that learners reported investing in ISLPs and the time that they spent on such activities. These contrasting findings might be indicative of differences between the learner populations (Lyrigkou studied slightly younger secondary school learners, aged 13 to 16, in Greece). They may also result from the different methodological approaches:

In this study, the participants self-reported their levels of linguistic engagement for each informal activity in the event-contingent diary, whereas Lyrigkou asked her learners to reflect more generally on their informal second language practices in a self-report questionnaire.

Altogether, the relationships discussed above show that the different engagement dimensions are strongly interconnected. That is, these variables seem to influence each other in various ways (see Figure 10.1), both positively (reinforcing) and negatively (inhibiting). At the participant-level, which focused on the pupils' general tendencies in engaging with ISLPs, affective, cognitive, and linguistic engagement were strongly interconnected, whereas behavioural engagement was mostly separate, except for a correlation between cognitive engagement and the quantity of ISLPs. This implies that the thoughts and emotions which learners generally associate with informal second language practices neither have a strong effect on, nor are strongly influenced by, the time that learners spend on ISLPs and the variety of activities in which they engage per week.

Meanwhile, at the activity-level, which captured relationships between these engagement dimensions within the same reported activity, all variables including behavioural engagement correlated significantly, if not always strongly, with one another. This suggests that whether or not a learner chose to continue to spend time on an informal activity was determined, to some extent, by the levels of focus it required, weighed against the enjoyment they derived from it. The fact that the patterns differed so much between the activity- and participant-level illustrates how multi-level modelling can help researchers make sense of dynamic constructs, such as learner engagement.

Throughout the above discussion, proficiency and task difficulty were mentioned as factors which the researcher thought may have affected the strength and direction of the relationships between the different engagement variables. The activity-level data (individual diary entries, spread over several days), could potentially provide further insight into how engagement can vary over time or according to the specific type of informal activity. Conducting these analyses was beyond the scope of the current study, but could be a worthwhile endeavour for further research. In order to gain a better understanding of ISLPs at a micro-level, future studies should also investigate how learners' engagement (their thoughts, attention, and emotions) fluctuates from moment to moment, within the same informal activity, and the variables which influence these changes.

Apart from the activity-level analyses discussed in this section, this thesis was mainly dedicated to studying engagement in ISLPs at the participant-level: participants' general tendencies in engagement, averaged across activities and time, which provide a more macro perspective. The next section discusses the results of the Latent Profile Analysis (LPA), which was used to explore whether the sample of students contained sub-groups of participants who showed distinct overall patterns of engagement in ISLPs.

10.1.4 Engagement Profiles

The multi-level analysis that provided the quantitative results discussed above was useful in revealing general patterns of engagement across the whole sample of participants. However, previous research has consistently reported large individual differences in informal second language practices (e.g. Kusyk, 2017b; Sockett, 2014b; Sundqvist, 2009, 2011; Sundqvist & Sylvén, 2014). In this study, too, there was high

within- and between-participant variability in all aspects of engagement. Therefore, the researcher decided to use 'person-centred' quantitative methods, in addition to the 'variable-focused' modelling techniques. Specifically, LPA was used to analyse the engagement data, which identified two sub-groups within the sample who showed distinct patterns of engagement in informal practices. The learners belonging to the first group spent significantly more time engaging in ISLPs than the second group, used a greater variety of resources, and showed higher levels of affective, but lower levels of linguistic engagement. There were no significant differences between the groups in terms of their average levels of cognitive engagement.

Interestingly, the two groups extracted from the LPA closely resembled two types of informal second language users which were observed in the qualitative data: First, a large number of the interviewed pupils described spending vast amounts of time engaging in many different informal practices, simply because the media which they enjoyed (the serials, social media content, games, etc. in which they were interested) 'happened' to be in English. Even if a German translation or dubbed version was available, many of these participants still opted for the English originals, because they perceived them as somehow more beautiful or expressive, or because they were dissatisfied with the quality of the German translations. These language perceptions suggest that these students were probably highly proficient in English. This was also illustrated in some of their claims that they did not perceive much of a difference between doing informal activities in German or English, and that they generally did not need to pay any conscious attention to the language in ISLPs. Altogether, the engagement patterns of these students seemed to match those of the first LPA group, in terms of the higher behavioural and affective and lower linguistic engagement.

By contrast, other focus group participants reported that they did not generally seek out, or enjoy, English-language informal practices, but might nevertheless engage in a particular activity if no German equivalent (e.g. a dubbed version of the same serial or film) was easily accessible to them. Often, these students stated explicitly that they enjoyed English-language practices less than they would the same activity in German, because of how much had to concentrate on understanding some of the language involved. This implies that these students were generally less proficient in English than the students described above. Their accounts seemed to match with the lower levels of behavioural and affective and higher levels of linguistic engagement of the second LPA group.

It is worth noting that, although both groups extracted in the LPA comprised an approximately equal number of students ($n_{Grp1} = 126$; $n_{Grp2} = 128$), most of the participants in Phase I could *not* be clearly matched to the two profiles described above, based on their interview responses. In part, this is because many participants did not describe all aspects of their engagement in ISLPs in such great detail. Yet, it could also indicate that there might be other engagement patterns, which were too uncommon to be identified in the quantitative analysis. LPA, being a 'bottom-up approach', is after all highly dependent on the specific sample data and the number of profiles the researcher chooses to extract. In turn, the number of profiles which appears to be optimal may, for example, be influenced by non-linearity in the relationships between the indicators (in this case, the five measures of engagement), or by other variables not included in the model (Bauer & Curran, 2004). Further research is therefore needed to validate the engagement patterns identified in this study.

A follow-up analysis was conducted to find out whether the two groups that were through LPA also differed with regard to the types of informal activities in which they primarily engaged. This revealed that the participants in the first group (higher behavioural and affective, lower linguistic engagement) were more than 3 times as likely as the participants in Group 2 to engage in mixed social media (which involved both reading texts and watching videos). Furthermore, the pupils in Group 1 also *spent more time* using mixed and video-based social media, as well as playing video games while chatting with other players, watching serials, listening to music (focused), and listening to spoken English.

Notably, the activities in which the first group engaged more often and for longer periods of time mainly involved listening. As mentioned previously (Section 10.1.3), there were some participants in the focus groups who specifically cited difficulties with listening comprehension as one of the main reasons for not engaging in informal second language practices more frequently. These pupils appeared to be less proficient in English than their peers who engaged in ISLPs more frequently. As some aspects of these pupils' general engagement in ISLPs matched that of the second LPA group (lower behavioural engagement and higher linguistic engagement), it seems that the lower popularity of listening activities among Group 2 might perhaps be attributable to differences in language proficiency. Indeed, further analyses found that participants with lower proficiency test scores were more likely to be part of the second LPA group (see Section 9.3.4).

The relationship between proficiency and engagement in ISLPs will be considered in more detail in the next section (10.2.). Thereafter, the discussion will turn to the relationships between engagement, language learning motivation, and attitudes towards language learning (Section 10.3.).

10.2 What are the Relationships Between General Language Proficiency and Engagement in Informal Second Language Practices?

One aim of the current study was to contribute to the growing body of research on the relationships between informal second language practices and L2 proficiency. This chapter will discuss the findings of both the qualitative and quantitative analyses, although the focus will be on the relationships between proficiency and engagement which emerged in the path models (Section 9.2). In short, while prior proficiency significantly predicted various aspects of engagement in ISLPs, engagement did not, in turn, significantly affect changes in the participants' general language proficiency over the course of this study.

In the next section (10.2.1), the researcher will discuss the implications of this finding, particularly in relation to the question whether ISLPs can promote the development of language proficiency. Thereafter (Section 10.2.2), the discussion will turn towards the relationships between specific aspects of engagement and proficiency that were found to be significant in this study.

10.2.1 Do Informal Second Language Practices Promote Proficiency?

This section discusses first the quantitative and then the qualitative findings relating to the question whether engagement in informal second language practices affects the development of L2 proficiency.

10.2.1.1 Quantitative Findings

In all of the tested models, language proficiency was found to have significant relationships with some of the variables measuring engagement in ISLPs. For example, prior proficiency alone explained 22% of the variance in the overall

engagement patterns (LPA groups), indicating a medium to strong effect. When looking at the separate engagement dimensions, as well as engagement in specific types of informal activities in which participants engaged, the effects were weaker but nonetheless statistically significant in many cases. Which specific aspects of engagement shared significant relationships with proficiency will be examined later (Section 10.2.2), but at this point it is worth mentioning that the two strongest direct relationships were between proficiency and the *quantity* and *variety* of ISLPs, both measures of behavioural engagement. Therein, the findings in this study align with those of the majority of previous ISLL research, which has found significant relationships between L2 proficiency and the frequency of engagement in informal activities or the time spent on them (e.g. De Wilde, Brysbaert, & Eyckmans, 2019.; Lyrigkou, 2016; Olsson, 2011; Sundqvist, 2009, 2011; Verspoor, de Bot, & Lowie, 2011; however, see Lai, Zhu & Gong, 2015; Lee, 2017, 2019; Lee & Dressman, 2017), as well as the variety of ISLPs (Lai et al., 2015; Lee, 2017, 2019; Lee & Dressman, 2017).

Despite that fact that these studies have overwhelmingly relied on correlational analyses (or, in some cases, simple and multiple regression analyses using data collected at only one time point), most researchers have interpreted these findings as showing that ISLPs can have a positive effect on the development of language proficiency. Previous attempts to investigate the directionality of this relationship have included research on ISLL in young learners, prior to receiving formal L2 instruction (De Wilde et al., 2019; Kuppens, 2007, 2010), as well as studies which compared the language skills of students with or without access to informal resources (M. H. Verspoor et al., 2011) and students with or without concurrent formal instruction (Cole, 2015). These studies have all provided some evidence for

the widely hypothesised positive effect of informal L2 practices on language proficiency. However, a large-scale, longitudinal study, conducted by Olsson and Sylvéén (2015), found no significant effects of the frequency and duration of informal language contact on the development of academic vocabulary over four years, even though the two variables had been significantly correlated at the first point of data collection.

Against this background, one of the important quantitative findings of the current study was that, in all of the analysed models, prior proficiency (T1 scores) significantly predicted some aspects of later engagement in informal second language practices, while engagement *did not*, in turn, significantly predict any changes in proficiency over time (T2 scores, controlling for T1). This does not mean that there was no relationship between engagement and posterior proficiency,²⁰ but none of the engagement measures predicted a significant amount of variance in the T2 scores beyond that already accounted for by the T1 measure. Because the sequence in which these variables were entered into the path model was specified by the researcher, the above finding does *not* provide directional evidence in and of itself, although the temporal order of the data does support this interpretation (for further discussion on this topic, see Wunsch, Russo, & Mouchart, 2010).

In part, the fact that engagement did not significantly affect general proficiency might be a reflection of the vast individual differences in informal second language practices that were the very the starting point for this study: Based on prior research, it seems

²⁰ In fact, post-hoc analyses revealed that the three engagement measures which were most strongly predicted by prior proficiency in the path models — the quantity and variety of ISL practices and the overall engagement patterns represented by the LPA group — shared nearly identical correlations with both the T1 and T2 proficiency scores (see Table F.2, Appendix F.2).

that whether ISLPs are successful at promoting learning will depend on a host of variables, including characteristics of the learners in question, the tools available to them, and the point at which they are in their own development (e.g. Kusyk, 2017b; Sockett, 2014b; Sundqvist & Sylvén, 2016). While the current study tried to control for a number of these variables through sampling restrictions and exclusions (see Section 3.3), it cannot be ruled out that there may have been significant variations between the sampled students in terms of other, non-controlled or unknown factors (e.g. affinity for technology, parental attitudes towards informal L2 use, etc.). Such individual differences, if they existed, may help to explain why this study not find a significant relationship between engagement in ISLPs and change in language proficiency over time. It is also possible that the pupils in this study may not have had enough (sufficiently engaging) informal second language contact within the time of the study for it to have had a significant impact on their general proficiency during this time.

Yet, it was not only engagement which failed to predict language development over the course of this study: In fact, proficiency at T1, which consistently explained 88% of the variance in the T2 scores, was the *only* significant predictor of the T2 scores in any of the analysed models. This was despite the fact that the motivational and attitudinal variables, which were also included in some of the path models (Section 9.5), are well established in the literature as predictors of language development (see Section 2.3). Thus, the total lack of significant predictors of language development, other than prior proficiency might indicate that (a) the period between the pre- and post-tests in this study (14 weeks) may not have been long enough for the participants to make significant progress in their general language development during this time;

or (b) the C-tests might not have been sensitive enough to measure any changes that did occur.

If the participants did not make measurable progress in their proficiency over the course of this study, this might be because they were quite advanced learners of English. A similar suggestion was made by Olsson and Sylvé (2015), whose longitudinal study, as stated previously, also failed to produce evidence for an effect of ISLPs on language development. The authors speculated. That this may have been due to the fact that their participants were Swedish upper secondary school students (aged 16–19) in English-medium education, who were highly proficient in English. Olsson and Sylvé suggest that the gains may have been greater for learners at lower levels of proficiency.

The participants in the current study also had relatively high starting proficiencies in English: In all three phases, the teachers estimated that their pupils who participated in this study had, on average, B1-B2 level competency (according to the CEFR, Council of Europe, 2001). Indeed, some pupils in the sampled population were likely to be even more advanced, as evidenced, for example, in (a) the detailed reflections they provided in the focus groups (Phase I) about the relative quality of original English-language media and German translations and (b) the fact that more difficult items had to be added to the C-tests, because a substantial proportion of participants in the pilot study (Phase II) scored at the upper end of the range.

Of course, even if the participants' language proficiency *did* develop significantly over the course of this study, the C-tests used to measure these abilities may not have been sensitive enough to pick up these changes. A paired-samples t-test revealed a significant difference between proficiency at T1 and T1 (sum scores weighted

according to the factor loadings in the measurement invariance models), with scores increasing, on average, by 7.3% (see Section 9.3.3). This tentatively suggests that there was indeed an increase in the proficiency of these participants over the course of the three months in which this study took place.²¹ However, this increase was slight enough that it must be questioned whether it occurred simply due to students' increased familiarity with the C-test format at T2 (before the start of this study, none of them had taken a C-test before). This is because researchers have proposed that score increases in the range of 5–6 percentage points, or 20–30% of the standard deviation (c. 3.8–5.7% in this study), should be expected when tests are repeated (Bogards & Raatz, 2002; Daller & Phelan, 2006).

As stated previously (Section 3.2), C-tests were employed as proficiency measures primarily because they have been argued to provide a valid measure of general language proficiency while also being time-efficient. Reducing the burden on the participants as much as possible was necessary to allow for the recruitment of enough school classes, given the time and resource constraints of this project, to collect a large enough data set for the sophisticated quantitative analyses. The data collection points, on the other hand, were spaced 14 weeks apart because this allowed for the evaluation of the relationship between ISLPs and language development over a longer period of time than in previous studies (e.g. Kusyk, 2017b; Sundqvist, 2009, 2011) and because it minimised the interference of the study with regular school operations (e.g. exam periods, holidays), again facilitating the recruitment of a larger sample.

²¹ It must be kept in mind that these mean scores should only be taken as a rough indication since the two C-test versions did not achieve strict measurement invariance, which statisticians consider essential to conducting valid mean comparisons (see Appendix E.4).

Nevertheless, given what has been discussed above, it must be recognised that employing only one measure of general language proficiency (instead of, for example, triangulating various measures of different aspects of language competence) and measuring language development over a relatively short period of time (three months) represent two core limitations of the current study. Thus, the hypothesis that ISLPs can contribute to language development should not be rejected without evidence from further longitudinal studies, which collect multiple measurements of various aspects of language proficiency and engagement in ISLPs, over longer periods of time.

10.2.1.2 Qualitative Findings

In contrast to the quantitative findings discussed above, some evidence for a positive impact of informal second language practices on proficiency, or at least the presence of such an effect in the learners' minds, could be found in the qualitative data collected in Phase I.

While the vast majority of the students who participated in the focus groups agreed that most of their current proficiency in English had come from school, all of them estimated that at least some of their language knowledge (between 10 and 40%) could be attributed to informal second language practices. This is much lower than in previous research, where some participants attributed as much as 70–100% of their current knowledge to informal practices (e.g. Cole, 2015; J. S. Lee, 2017). However, the participants in the current study also showed high variability in terms of their involvement in ISLPs, whereas, for example, the case-studies described by Lee (2017) and the 'fully autonomous self-instructed learners' in Cole (2015) were selected specifically as representatives of highly engaged informal learners. The 'classroom-

trained learners' in Cole's study, which were presumed to be more representative of the wider learner population in this context seemed to assign less importance to informal L2 contact. Only 27.5% believed 'that an informal resource [had been] most important for their language learning' (p. 182).

It is important to consider that whether or not individual participants in the current study perceived ISLPs as effective depended heavily on how they perceived *proficiency*: Those pupils who stressed, elsewhere in the interviews, that English was important to them because they wanted to use it to communicate with others were usually more positive about the effectiveness of informal learning (implying that they thought ISLPs contributed to improving their communicative competency) than those who judged their language proficiency in reference to standards in formal education (e.g. exam grades, grammatical accuracy). In fact, the specific (German) words used to discuss learning gains appeared to have vast effects on how the participants perceived the effectiveness of ISLL: The overwhelming majority of pupils did not think of their informal practices as activities for 'learning' (in German: 'lernen'), but did agree that they contributed to 'improving' ('verbessern') and practicing ('üben') their English skills.

The main benefit of engaging in ISLPS, as perceived by the interviewed students, was that they were able to expand and individualise their vocabularies through choosing the topics with which they wanted to engage. This in accordance with earlier research, which has portrayed vocabulary acquisition as one of the main benefits of ISLPs (J. S. Lee, 2017; Sockett & Toffoli, 2012; Sundqvist, 2009, 2011). By contrast, the interviewed students felt that they did not learn a lot of new vocabulary in school, and

that the words they were taught (especially in their more recent, advanced classes) were not often personally relevant to them.

The participants furthermore suggested that there were great differences between the various types of informal activities regarding what could be learned from engaging them. For example, speaking in English and reading books were most often mentioned as examples of 'effective' activities, although the pupils recognised that these were not ones in which they tended to engage. A few participants did suggest, however, that their reading comprehension and speed had been positively impacted by frequently reading English-language texts on social media and the internet in general. The activities in which the majority of these students did engage frequently (including watching serials, listening to music, and accessing multimodal content [text and videos] through social media) primarily involved receptive language processing. Therefore, the participants thought that these practices would not substantially affect their *productive* language skills (specifically fluency in writing and speaking), to which they assigned the most value. Yet, similarly to a group of French university students interviewed by Sockett and Toffoli (2012), many pupils in the current study did think that watching online videos or serials had improved their listening comprehension and helped them learn to understand various accents.

In contrast to the perceived benefits of informal second language practices, the participants overwhelmingly credited their school classes with being the source of all of their grammatical and syntactical knowledge and forcing them to practice writing and speaking in English (which most of them did not do in their free time). Furthermore, they acknowledged that they had been able to reach a 'basic level' of English through formal instruction, which then allowed them to use the language in

their free time, to pursue their personal interests or to access English-language media for entertainment.

On the one hand, this supports the hypothesis that ISLL is effective only when students have already acquired a good baseline knowledge of English, although there are studies which have provided evidence of successful informal learning of English in very young, less proficient learners, some of whom had not yet received any formal L2 instruction (De Wilde et al., 2019; Kuppens, 2007, 2010; Sundqvist & Sylvén, 2014). On the other hand, these findings suggest that, instead of debating the relative effectiveness of informal and formal learning, it might be more valuable to view both sets of practices as complimentary parts of a well-balanced *language learning diet*. Furthermore, the vast amounts of time that students have been shown to spend on informal practices in this and other studies must also be considered, especially compared to the few hours of language instruction per week which they receive in schools. Thus, even if future research should show ISLL to be relatively less efficient than formal instruction, understanding, promoting, and improving students' second language practices outside of the classroom will remain a worthwhile endeavour for researchers, teachers, and policy makers.

The findings that are discussed in the next section contribute to this endeavour by providing further insight into how the various aspects of engagement in informal activities that were measured in this study were related to language proficiency. These relationships will be described and interpreted as they emerged from the models, that is as showing that engagement is, to some extent, determined by learners' prior proficiency. However, given the limitations of the longitudinal measures in this study as discussed in Section 10.2.1.1, the hypothesis that ISLPs can

contribute to language development should not be rejected on the basis of the current study, without further evidence. Future research may well show that the various dimensions of engagement can predict variations in proficiency — and thus explain some of the individual and group differences discovered in previous ISLL research (see Chapter 2) — or that these factors mutually affect one another over time, an option which seems not just theoretically plausible, but indeed rather likely.

10.2.2 Relationships Between Proficiency and Various Aspects of Engagement

This section focuses on explaining and interpreting the results of the quantitative analyses which statistically modelled the relationships between proficiency and the various measures of engagement in ISLPs: (1) the general engagement patterns, represented by the LPA groups; (2) the five aspects of engagement (quantity and variety of informal practices [behavioural engagement], affective, cognitive, and linguistic engagement); and (3) the time spent on specific types of activities. Wherever possible, the findings are further contextualised using insights from the qualitative data.

10.2.2.1 Engagement Profiles

To begin with, prior proficiency was a significant predictor of the participants' LPA group membership, that is, of their overall patterns of engagement in informal second language practices (including behavioural, affective, cognitive, and linguistic engagement). As noted above (Section 10.2.1.1), proficiency at T1 explained 22% of the variance in the binary LPA group variable, providing strong evidence for the hypothesised relationship between engagement in ISLPs and proficiency. Specifically, participants who achieved higher C-test scores at T1 were more likely to be part of

the first LPA group, who spent significantly more time on a more diverse range of ISLPs (higher behavioural engagement), showed more interest in informal practices and derived greater enjoyment from them (higher affective engagement), but generally focused less on the language features they encountered (lower linguistic engagement). The relationships between each of these dimensions of engagement and proficiency that emerged from the remaining analyses and will be discussed in turn below.

10.2.2.2 Behavioural Engagement: Quantity and Variety

Proficiency at T1 had significant, positive effects on the time that learners spent on ISLPs (quantity) and the range of activities in which they engaged (variety), both aspects of behavioural engagement. Prior proficiency alone explained 17% and 12% in the quantity and variety of ISLPs respectively. These results are consistent with prior studies, the majority of which have found positive relationships between the quantity or frequency of engagement in ISLPs and proficiency with similar effect sizes to the current study (e.g. Cole, 2015; De Wilde et al., 2019; Lyrigkou, 2016; Olsson, 2011; Sundqvist, 2009, 2011). As mentioned previously, most researchers have interpreted this as evidence for a positive effect of informal second language practices on proficiency. Yet, as these studies have mostly employed correlational methods, their findings equally support the conclusion that learners with greater proficiency are more likely to engage in informal practices, which was shown in the current study.

Interestingly, among the only pieces of research which did *not* find a positive relationship between the quantity of ISLPs and proficiency are two studies which also measured the variety of activities, as in the current thesis. Both Lai, Zhu, and Gong (2015) and Lee (2017, 2019) found significant, moderate to strong correlations

between the number of different English-language activities in which learners engaged (variety of ISLPs) and various measures of proficiency. At the same time, they recorded no significant correlations between proficiency and the duration of the learners' out-of-school practices (quantity of ISLPs). There are two features setting these studies apart from the current thesis and other ISLL research, which may offer explanations for the divergence in findings.

First, in both studies, the out-of-school activities that were considered included both 'meaning-focused activities' (equivalent to informal second language practices in the current studies) as well as 'form-focused activities' (*non-formal* activities such as grammar exercises, reviewing textbooks, and attending tutorial classes). While the participants spent more time on meaning-focused activities, form-focused practices were also prominent in these samples, which is in contrast to the results of the current study and other ISLL research (see e.g. Kussyk, 2017b; Sockett, 2014b; Toffoli & Sockett, 2010). Lai, Zhu, and Gong (2015) separately calculated correlations between meaning-focused activities and proficiency, and still only the variety, but not the quantity of these practices emerged as significant. However, it seems possible that the learners' overall stronger participation in non-formal practices in this and Lee's (2017, 2019) study might be a reason why out-of-school practices did not seem to contribute as much to the participants' language development as for learners in other studies, for whom ISLPs were the primary means of L2 contact outside of school.

Second, while the majority of ISLL research has been conducted in Europe,²² the studies by Lai and associates (2015) and Lee (2017, 2019) were both set in East

²² However, see e.g. Cole's (2015) research in Brazil, Lamb and Arisandy's (2019) work in Indonesia.

Asian contexts. The former focused on the practices of secondary school students learning English in China, while the latter involved adult EFL learners at three South Korean universities. Thus, it must also be considered whether the lack of a relationship between the quantity of ISLPs and proficiency in these studies might, in some part, be due to differences in the socio-cultural and educational contexts. As noted repeatedly throughout this thesis, the field of ISLL would greatly benefit from larger, cross-national and longitudinal studies which would enable the comparison of patterns in informal learning practices and language development, for example across learners of different ages and from different cultural and educational contexts.

10.2.2.3 Behavioural Engagement: Types of Activities

Besides the quantity and variety of the participants' informal second language practices, this study also explored the relationships between proficiency and the types of activities in which the pupils engaged. Prior proficiency successfully predicted how likely participants were to engage in six out of the thirteen informal activity categories: Specifically, there were significant small to medium-sized effects of T1 proficiency on the *likelihood* of watching films and serials, listening to spoken English, video-based and mixed social media use (i.e. involving video- and text-based activities), and focused music listening. All of these effects were positive, meaning that more proficient students were more likely to engage in these activities. It is noteworthy that this list includes almost all practices which mainly involved listening, except for *listening to music in the background*.

The finding that proficiency primarily predicted engagement in listening activities can be linked to the observation, noted several times throughout the preceding discussion, that some of the students interviewed in Phase I reported

having difficulties with listening comprehension, which they cited as one of the main reasons for not engaging in ISLPs more. The researcher had speculated that these were most likely pupils with lower proficiency in English, thus demonstrating an effect of proficiency on the choice not to engage in particular informal activities. The above quantitative findings, although they stem from a different sample of participants within the same population, support this conclusion.

When it came to predicting the *amount of time* that participants spent on individual activities, only four significant effects emerged: Prior proficiency predicted the time spent reading and on all three categories of social media (text-based, video-based, and mixed), with small to medium effect sizes. Further simple linear regressions were conducted to investigate whether proficiency specifically affected the time that participants spent on particular social media platforms. Only one significant effect was found, which indicated that students with higher language proficiency spent significantly more time on YouTube. Although this effect was very small (proficiency explained 2% of the variance in time spent watching YouTube videos), it helps to explain why more proficient students were found to spend significantly more time on both video-based and mixed social media, as YouTube accounted for 79.5% and 43.8% of the entries in these categories, respectively.

Once again, this finding can also be linked to the particular difficulties which some of the students experienced with listening comprehension. In particular, a number of the participants in the focus groups, who the researcher speculated might be less proficient than their peers, mentioned struggling with the fact that YouTubers often spoke very quickly, which they found difficult to concentrate on and to understand. However, a similar number of interviewees also mentioned struggling with the rate of speech and with different accents when watching serials and films,

while no significant effect of proficiency was found on the time the participants in Phase III spent on these activities. A possible explanation might be that all learners, regardless of their proficiency, spend more time on serials and films because these are both longer (i.e. the length of an episode or a movie as compared to one online video) and narratively more immersive than YouTube videos, as viewers become emotionally invested in the characters' stories.

Besides watching online videos and serials, reading was also frequently mentioned in the focus groups as an especially difficult activity, particularly reading books in English. Again, this might explain why prior proficiency was linked to the time spent on reading activities, which included reading books, magazines, and online articles. Somewhat surprisingly, this effect also seemed to transfer to text-based social media use, which the researcher would have expected to be primarily determined by personal interests and social motives. Nevertheless, proficiency at T1 explained almost as much of the variance in the time spent on text-based social media as on other reading (13% vs. 14%).

The link between language proficiency and social media use in this study is in accordance with the findings of several prior studies of ISLL (Cole, 2015; De Wilde et al., 2019; Lyriqkou, 2016). Sundqvist (2009, 2011) also found a strong relationship between the time her participants spent on reading activities ('surfing the web' and reading books, newspapers, and magazines) and their vocabulary and oral proficiency test scores. However, these studies have not generally reported strong links between listening activities and proficiency, as were found in the current data: Sundqvist (2009, 2011), for example, found that the time her participants spent listening to music and watching TV and films were relatively much more weakly (although still

significantly) related to proficiency than playing video games and the aforementioned reading activities. Similarly, Cole (2015), found many significant differences regarding engagement in various activity types between the highly proficient, ‘fully autonomous’ informal learners and less proficient, ‘classroom-trained’ learners, but reported that these groups did *not* differ significantly in their engagement with music, television, and film — the only listening activities examined in his study.

Finally, De Wilde, Brysbaert & Eyckmans’ (2019) reported that speaking, using social media and playing video games in English, but none of the listening activities in their study, significantly predicted proficiency (including measures of reading, writing, listening, speaking, and vocabulary). Overall, the finding that listening activities were more strongly linked to proficiency in this than in other studies, together with the frequent mentions of difficulties with listening comprehension in the focus groups, suggests that listening may be relatively underdeveloped in this particular population of learners as compared to the other language skills.

10.2.2.4 Affective Engagement

Based on the findings from the focus group interviews, where enjoyment and interest emerged as the main determinants of participants’ willingness to participate in informal activities, the researcher hypothesised that participants’ affective engagement would be affected by their proficiency. That is, it was expected that more proficient participants would derive more enjoyment from informal activities, because they would be less likely to feel frustrated from encountering issues with comprehension. Yet, no significant direct links were found between proficiency and participants’ interest in and enjoyment of the ISLPs (affective engagement) in any of the path models.

To understand this finding, it is helpful to consider the timeline of the available data on affective engagement: The qualitative data revealed that enjoyment and interest mainly seemed to be important at the point where pupils decided whether or not to engage in an activity, or if they were considering whether or not they should continue with an activity in which they were already participating. More precisely, then, this choice was found to be affected by the *potential for affective engagement* that students saw in an informal activity (how much enjoyment they expected to derive from it). It seems likely that this perception, in turn, would be influenced by their *past* affective engagement: the enjoyment they had derived previously from engaging in the same kinds of informal activities. However, the quantitative data on affective engagement consisted of retrospective self-reports of participants' levels of enjoyment of the ISLPs. Thus, the quantitative findings merely suggest that, once learners *were already engaged* in an informal activity, their actual enjoyment was not significantly linked to their proficiency. Meanwhile, it is still possible, and indeed it seems likely, that learners' language skills would affect their decision about whether or not to engage in ISLPs in the first place, via their potential for affective engagement.

Whether or not this is found to be true, the above explanation illustrates the complexity inherent in the variables in this study and the relationships between them. As discussed in the literature review (Section 2.2), engagement is a multifaceted construct, conceptualised both as relatively stable (reflecting individual tendencies) and dynamic across time (varying in relation to changes in both the context and learner characteristics). The current study aimed to capture some of this complexity by separately modelling the relationships between the different dimensions of engagement at both the activity- and participant-level. However, engagement was still treated as a static construct; the study did not, for example, model the changes in the

engagement of the same participants across different activities or time. Furthermore, only the relationships between the participant-level engagement scores (i.e. per-learner averages of separate instances of engagement in ISLPs) and other learner variables (proficiency, motivation, and attitudes) were explored statistically. Further research is clearly needed to understand the complex nature of engagement in ISLPs, including the micro- and macro-level roles of the different engagement dimensions and how they change in relation to each other and other variables over time.

10.2.2.5 Cognitive and Linguistic Engagement

While a positive relationship was found between proficiency and behavioural engagement (Section 10.2.2.1), both cognitive and linguistic engagement were *negatively* predicted by prior language skills. That is, less proficient students were likely to be more concentrated on the ISLPs in which they engaged, and to pay more conscious attention to the language features they encountered. This conforms to the researcher's earlier speculations, that the focus group participants who described having to focus more intently during informal activities in order to understand the meaning of the language they encountered were probably less proficient than their peers, for whom reading or listening to English felt similar to German (their first language).

The effects of prior proficiency on both cognitive and linguistic engagement were about equal in strength. As suggested previously (Section 9.5.1), the fact that both of these negative, direct effects reached statistical significance only when modelled concomitantly with motivation seems to be explained by them being cancelled out in the first model by positive, indirect effects via motivation (see Section 10.3 below).

One aim in including measures of cognitive and linguistic engagement in the current study was to further investigate Cole's (2015) finding that the amount of 'mental effort' invested in informal activities, together with differences in motivation and attitudes towards ISLL, seemed to explain differences in learning outcomes associated with ISLPs. Specifically, Cole suggested that the 'fully autonomous self-instructed learners' in his study might have been more proficient than the 'classroom-trained learners' because they were (a) more motivated to master the English language and considered informal practices as valuable learning opportunities (rather than merely as leisure activities), thus (b) investing more effort into processing the meaning and language in the informal activities. In the current study, evidence was found in support of Cole's first finding: some aspects of motivation and attitudes towards ISLL did indeed have significant effects on cognitive and linguistic engagement (see Section 10.3). However, the findings regarding the link between cognitive effort/focus on form and language proficiency are less clear.

Even though cognitive and linguistic engagement did not significantly predict language development over the course of this study, a significant relationship did emerge between engagement and prior proficiency. However, as discussed previously (Section 10.2.1.1), firm conclusions cannot be drawn about the directionality of these relationships in the current study. Nevertheless, the fact that this relationship was *negative* runs contrary to Cole's (2015) finding that more proficient learners seemed to invest more mental effort while engaging in informal second language practices (see also Lai et al., 2015; Lyrigkou, 2016).

Therein, the findings of the current study also speak to a broader issue, which has been the subject of much theoretical debate and empirical research in SLA:

the role of attention and awareness in language learning. It has been argued extensively that at least some conscious attention to the linguistic features encountered in the input is necessary for learning to occur (cf. Noticing Hypothesis: Schmidt, 1990). Nevertheless, many studies also show that *incidental learning* (e.g. Hulstijn, 2001; Nation, 2001), especially lexical learning, can occur when learners are engaged in an activity where their intention is not to learn, or if they are mainly focused on understanding the meaning (e.g. see review in Leow & Zamora, 2017).

If we choose to accept the hypothesis that ISLPs can promote language development, as is widely presumed in the ISLL literature, the comparative strength of the links that were found between various dimensions of engagement and proficiency in this study imply that the quantity of language to which individuals are exposed and the variety of activities in which they engage are the most important factors in informal learning, rather than the amount of conscious, strategic attention paid to the linguistic features. Therein, the current study provides some evidence for the claim that ISLPs mainly involves incidental learning processes, i.e. that it is effective mainly because of the extensive opportunities for interesting and personally meaningful second language use that they offer (Benson & Chik, 2010; Sykes, Oksoz, & Thorne, 2008; Toffoli & Sockett, 2010).

Nevertheless, as noted above (Section 10.1.2), some of the pupils interviewed in Phase I stated that if they encountered issues with comprehension while engaging in informal activities, they would sometimes focus on a particular linguistic feature in order to increase their understanding of the overall meaning. This suggests that, although learners may broadly conceive of informal activities as mainly for entertainment, with very little overall linguistic engagement, they may nonetheless pay conscious attention to decoding and learning new linguistic forms momentarily.

The retrospective self-report questions used to measure engagement in this study prompted students to estimate how much they focused on features of the language during the activity as a whole. Therein, this instrument may not have been sensitive enough to detect the kind of micro-level linguistic engagement to which the qualitative data alluded. This suggests that further research is needed which investigates engagement at a micro-level, capturing changes which may occur from moment to moment. For example, in order to gain a better understanding of the role of 'noticing' in ISLL, screen capture and eye-tracking methods may be combined to investigate where learners place their attention while engaging in particular informal second language activities (e.g. Godfroid, Housen, & Boers, 2010; Godfroid & Winke, 2015).

10.3 What are the Relationships Between Motivation and Attitudes Towards

Language Learning and Engagement in Informal Second Language Practices?

In this section of the discussion, the focus is on the relationships between engagement in informal second language learning and three groups of motivational and attitudinal variables measured in this study: (a) the *language learning orientations*, representing learners' perceptions of the English as a second language and why it would be useful for them to know it; (2) the learners' current and future L2 self-concepts, and the difference between them, which reflect the pupils' perceptions of themselves as language learners and users; and (3) the participants' perceptions of the learning context, specifically their attitudes towards informal and formal learning. While the discussion will draw on findings from both the qualitative and quantitative analyses, the primary focus, as before, will be on the relationships which emerged as significant in the path models.

All of the engagement variables were significantly predicted by some of the motivational factors at T1. In turn, significant changes in only three of the motivation variables were predicted by engagement. As with the lack of significant effects of engagement on language proficiency (discussed in Section 10.2), caution must be exercised when interpreting these findings as evidence for a directional relationship. Again, we must consider the possibility that the time between the first and second data collection (14 weeks) may have been too brief for substantial and measurable changes in some of the motivation constructs to occur. Yet, while proficiency at T2 lacked any predictors besides prior proficiency, the changes in the motivational and attitudinal variables were predicted, as expected, by both proficiency and other motivational factors at T1 (see Appendix I.1). This lends validity to the measures of changes in motivation and attitudes over the course of the study and, together with the temporal order of the data, tentatively supports the conclusion that motivation and attitudes may have greater effects on engagement in ISLPs than vice-versa.

10.3.1 Language Learning Orientations

Of the five language learning orientations (LLOs) measured in this study, only three emerged as significant predictors of engagement in ISLPs:

1. *Knowledge orientation*: wanting to know English to pursue personal interests, access information, and help in acquiring new knowledge
2. *Intrinsic motivation*: wanting to know English because of an inherent affinity for the language and finding the learning process enjoyable
3. *Instrumental obligation*: feeling pressured to learn the language, for example by teachers, parents, or school more generally.

For the most part, these motivational orientations had positive effects on engagement. They primarily affected the learner-internal aspects (affective, cognitive, and linguistic engagement), although knowledge orientation and instrumental obligation also shared significant relationships with the quantity and variety of ISLPs respectively (behavioural engagement). The effects of each of these LLOs on engagement will be discussed in more detail below.

Knowledge orientation

The relevance of knowledge orientation for behavioural engagement had already emerged in the qualitative data collected in Phase I. The interviewed participants often stressed that, to them, the most important criterion for choosing informal activities was that they must be ‘interesting’ (in German: ‘interessant’). In most cases, this meant that a specific piece of English-language popular culture (a song, serials, video game, etc.) had piqued their curiosity enough to try it. Yet, many pupils also spoke about being ‘interested’ in particular subjects (e.g. beauty, fashion, science, sports, politics), which they were able to engage with and learn more about using English-language resources, such as online videos and texts, and even social media.

It is unsurprising then that the time participants spent on ISLPs was significantly predicted by knowledge orientation, which measured the extent to which the learners perceived the English language as a tool that helped them pursue personal interests, access information, and acquire new knowledge. A similar finding emerged in Lyriqkou’s (2016) study, where the learners’ frequency of informal contact was significantly correlated to how much they agreed with the statement ‘I think that contact with English through informal sources helps me to access the information and entertainment that is personally important to me’ (p. 108). In the current study, a high

knowledge orientation also positively predicted affective and cognitive engagement. That is, students who felt that English empowered them to pursue their personal interests and educate themselves were more emotionally involved and focused when engaging in ISLPs.

In interpreting all of these findings, the special status of English (as compared to other foreign languages) in the lives of these students must be taken into account. The fact that the informal use of English was so closely tied to their interests speaks, in part, to the strong global influence of English-language (particularly American) popular culture (e.g. Elsworth, 2015), as well as to the role of English as a lingua franca (Seidlhofer, 2005), that is, as a primary means of information dissemination and communication among both native-and non-native speakers of the language, particularly in online discourse.²³ Therefore, all results from the present study, but particularly the findings regarding the role of knowledge orientation, must be validated through further research with other populations and target languages.

Intrinsic motivation

Based on the qualitative findings, the researcher also expected to find a significant relationship between intrinsic motivation and the quantity of ISLPs: Some interviewees reported that they engaged in informal practices in English because they had an intrinsic preference for the language, perceiving it, in their own words, as more expressive, beautiful, or pleasant than German. Others had an interest in learning and using languages in general and saw informal second language use as a fun challenge.

²³ English is the most prominent language on the internet, with an estimated 54% of all online content being written in English ("Usage of content languages for websites," 2019). This is even more so the case on social media and content sharing websites: For example, in 2016, 95% of the contents uploaded to YouTube per day was in English (Turek, 2016).

However, in the quantitative models, intrinsic motivation did not emerge as a significant predictor of behavioural engagement. This may mean that the highly intrinsically motivated students in the interviews were exceptional cases, so that the effect on behavioural engagement was not strong enough overall to reach statistical significance in the full sample.

Instead, intrinsic motivation did significantly predict all three of the learner-internal engagement factors: affective, cognitive, and linguistic engagement. Students with high intrinsic motivation wanted to know English because they liked the language and found the learning process enjoyable. The quantitative findings show, then, that this general passion for the language translated directly to the learners' cognitive and emotional states while engaging in ISLPs. That is, they derived more enjoyment from informal activities (affective engagement), approached them with more focus (cognitive engagement), and paid closer attention to the language they encountered (linguistic engagement).

Instrumental obligation

Instrumental obligation was perhaps the most interesting LLO because it showed *positive* effects on affective, cognitive, and linguistic engagement while also being a *negative* predictor of the variety of ISLPs. An effect of instrumental obligation on the quantity of ISLPs had been present in the earlier models (Section 9.4.1.1) but became non-significant after the addition of proficiency (Section 9.5.1). Thus, it appears that instrumental obligation did not directly affect the time spent on ISLPs, but rather that proficiency negatively affected both instrumental obligation and the quantity of ISLPs: Less proficient students were more likely to feel that they were only learning English because they had to, and they were also less likely to engage with the language

outside of school. The lack of a direct relationship between instrumental obligation and the quantity of ISLPs also resembles an earlier finding by Lamb and Arisandy (2019), who discovered no significant correlation between the ought-to self ('measuring how far participants thought that significant others in their lives believed they should learn English', p. 9) and their participants' likelihood of engaging in 'entertainment-oriented' activities outside of school.

It is unclear at this moment why the other measure of behavioural engagement in the current study, the variety of ISLPs, would behave differently in this regard, being in fact directly predicted by instrumental obligation even after the inclusion of proficiency in the same model. Yet, the finding that these two aspects of behavioural engagement should behave differently from one another is not unique to the current study: Some prior ISLL studies have recorded positive relationships between proficiency and the variety, but not the quantity of informal practices, which led the researchers to conclude that the type of activities in which the learners engage might be of greater importance to developing language proficiency than the time spent on them (Lai et al., 2015; J. S. Lee, 2017; J. S. Lee & Dressman, 2017).

The positive correlation between instrumental obligation and affective engagement implies that students who felt more pressured to learn English actually found informal activities to be more enjoyable and interesting. Surprisingly, this effect was even stronger than the effect of the knowledge orientation on affective engagement. Perhaps, the greater enjoyment of participants with strong instrumental orientation can be attributed to the pupils deriving enjoyment from discovering ways of using the language which felt interesting and meaningful to them, rather than like something that they were obligated to do. It may also be possible that, for pupils who feel that learning English was a bothersome obligation, the 'threshold of interest' that

an informal L2 activity had to cross in their minds to be worth doing was higher (see Section 10.1.2). This would explain why these students recorded comparably higher levels of affective engagement for activities in which they did, indeed, choose to participate.

The higher levels of cognitive and linguistic engagement associated with the instrumental obligation, on the other hand, could be indicative of the students trying to apply formal learning strategies to informal activities. Such transfer of learning strategies acquired at school was observed, for example, in recent studies on autonomous adult L2 learners in Hong Kong (Chik, 2014, 2018; Lai, 2015). In other words, if a student generally felt pressured to improve their English skills, they may have been more inclined to think of ISLPs as learning activities, leading them to concentrate more on the activity and to pay more attention to language features than other students, who perceived informal activities primarily as opportunities to pursue personal interests and entertainment. This explanation is based on Cole's (2015) suggestion that individuals who consider informal practices as valuable learning opportunities, rather than merely as leisure activities, invest more effort into processing their meaning and learning new linguistic features they encounter.

Lastly, instrumental obligation was one of the only variables in which changes over the course of the study were significantly predicted by engagement. In particular, being generally more focused while engaging in informal activities (higher cognitive engagement) was predictive of a *decrease* in the sense of external pressure that participants felt to learn English (instrumental obligation). Similar to the positive relationship between affective engagement and instrumental obligation, this implies that a potential benefit of engaging in ISLPs is showing pupils ways in which knowing English can be personally valuable and meaningful to them. The finding that

specifically cognitive engagement predicted a decrease in instrumental obligation implies that a focus on meaning in informal language use is beneficial for strengthening the students' sense of the inherent value of the language.

Against this background, the finding that greater affective engagement predicted an *increase* in instrumental obligation seems counter-intuitive at first. However, it helps to view this result through the lens of self-determination theory (SDT; R. M. Ryan & Deci, 2000), which places extrinsic and intrinsic motivation on a continuum and stresses the role of the 'perceived locus of causality' in internalising goals (Deci & Ryan, 1985, p. 68). In SDT terms, having positive experiences of informal L2 use (high affective engagement) may lead students to recognise that there are benefits to knowing English, which may cause them to value English as a tool to pursue their own goals and interests (internal locus of causality; cf. knowledge orientation). Yet, if pupils feel like they cannot pursue personal interests or goals *without* knowing English, ISLPs could also contribute to students feeling even more pressured from the outside to learn the language (external locus of causality). The second of these outcomes, then, might be reflected in the effect of affective engagement on instrumental obligation in this study: In other words, enjoying informal activities strengthened pupils' feelings that they *ought to* learn the language better (instrumental obligation). Yet, it seems possible that, over time, this motivation may become more internalised in the longer term, if learners start to recognise informal second language activities as more personally valuable to them.

Finally, it is worth considering the two LLOs which did not show any significant relationships to engagement in ISLL: *international orientation* and *instrumental benefits*. The former was defined as wanting to know English to be able to interact

with people from other cultural and language backgrounds, and its irrelevance to engagement in ISLPs is not altogether surprising when considering the types of activities in which the participants in this study engaged most often. While the opportunities for active participation, meaningful interaction and collaboration with other L2 learners are among the most highlighted affordances of the internet (e.g. Greenhow, Robelia, & Hughes, 2009; Mills, 2011; Thorne, Black, & Sykes, 2009), most of the participants in this study rarely or never interacted with others informally in English unless they had personal connections to English speakers, for example, via family links or through having attended an international exchange program. This reflects trends in prior ISLL research (e.g. Lai et al., 2018; Sockett, 2014c; Sundqvist, 2009, 2011) and helps to explain the lack of a relationship between the international orientation and engagement in ISLPs in this study.

Lastly, the instrumental benefits orientation, measuring the extent to which the participants were motivated by the advantages they believed English offered them in terms of their future career and education, was also not significantly related to any aspect of engagement. This may be a result of the relatively high proficiency of the learners in this study, as it has been suggested that this orientation primarily predicts students' effort at the early stages of language learning (X. Liu, 2016). Another possible explanation can be found by taking a closer look at the way in which the participants in this study conceptualised language in formal and informal contexts (see also Section 10.3.3): The vast majority of the participants interviewed in Phase I agreed that they engaged in ISLPs solely for the purpose of entertainment, not for learning. They also largely felt that most of their current English proficiency had come from school, not ISLPs, and that, informal activities mainly contributed to their knowledge of colloquial English (as opposed to more formal *school English*

[in German: ‘Schulenglisch’]]. Thus, if the participants did not think that informal practices were relevant to their acquisition of the type of English that would be useful to them in their later studies or career, this may explain why the instrumental benefits orientation did not predict engagement in ISLPs, and vice versa.

10.3.2 Second Language Motivational Selves

Inspired by Dörnyei’s (2005) theory of the Second Language Motivational Self-System (see Section 2.3), the participants in this study were asked about their current self-perceptions (current L2 self) as well as the extent to which they imagined themselves as successful future users of the language (ideal future L2 self). Given that Dörnyei’s theory hinges on the idea that learners’ behaviour is motivated by a desire to reduce the distance between the current and ideal future self, *motivational distance* was also included as a separate variable in the path models.

Current L2 self

A number of previous ISLL studies have investigated variables similar to the current L2 self and generally found them to have positive relationships with behavioural engagement in informal second language practices. Sundqvist (2009, 2011), for example, found significant positive correlations between self-efficacy beliefs of secondary school learners in Sweden and the quantity of their informal contact with English. Both Lai, Zhu, and Gong (2015) and Lee (2019) discovered similar-sized correlations between the confidence learners had in their ability to learn English and the variety of informal activities in which they engaged. In Lee’s (2019) study, confidence also significantly correlated with the time learners spent on ISLPs, whereas this was not the case in Lai, Zhu, and Gong’s (2015) study. However, none of

these studies looked at the relationship between ISLPs and participants' self-perceptions while also controlling for their actual proficiency, which turned out to be essential in the current study.

When the effects of motivation and proficiency were modelled separately, the current self at T1 had positive effects on the overall patterns of engagement (as represented by the LPA groups; see Section 9.4.2) as well as both measures of behavioural engagement (Section 9.4.1.2). This implied that pupils who judged themselves to be more successful L2 users currently were likely to spend more time engaging in a greater variety of informal activities. Moreover, in this model, the variety of ISLPs was also found to have a very small, but significant positive effect on the changes in participants' current self over the course of the study. However, all of these relationships were no longer significant after controlling for proficiency in the same model (Section 9.5.1). This suggests that actual proficiency may play a more important role in ISLPs than the pupils' perceptions of their own language abilities.

Future L2 self

The future L2 self, on the other hand, significantly predicted the variety of informal activities in which the participants engaged, even after controlling for actual proficiency at T1. In fact, the two variables mutually affected one another, as variety also had a small but significant effect on changes in the participants' ideal future self-concept over the course of the study. This means that pupils who had a strong vision of themselves as competent future L2 users were more likely to use a greater variety of informal resources, and that, in turn, using English in a variety of ways beyond school helped to strengthen the learners' desire to become successful language users.

To this researcher's knowledge, only one prior study has investigated the role of the ideal L2 self in ISLL: Lamb and Arisandy (2019) discovered significant positive correlations between the strength of the ideal L2 self of university-level EFL learners in Indonesia and their engagement in L2 activities beyond the classroom. Data on informal activities were gathered in their study by asking participants to rate how likely they were to engage in each of a list of 27 different activities, using a Likert scale that ranged from 'not at all true of me' to 'very true of me' (p. 8). In a way, this measure thus concurrently evaluates both the frequency and variety of engagement in ISLPs. Therefore, Lamb and Arisandy's findings seem to be in accordance with the current study, providing partial evidence for a positive relationship between the ideal future self and the variety of ISL practices. In future, researchers may wish to explore whether this relationship arises due to a conscious learner choice, that is, whether pupils purposefully choose to engage with English in diverse ways to strengthen their language skills. Furthermore, the above findings also suggest that encouraging learners to try different types of informal second language practices might help to increase their motivation for language learning.

Motivational distance

Finally, the motivational distance shared significant relationships with affective, cognitive, and linguistic engagement, none of which emerged as significant when the current and future L2 self were modelled separately. This must mean that these engagement dimensions affected both aspects of the L2 self-concept concurrently: together, a non-significant increase in the strength of the current L2 self with a simultaneous, non-significant decrease in the ideal future self would have resulted in a significant decrease in the motivational distance, and vice versa. The presence of

these effects thus illustrates how considering the distance between the current and future L2 self can lead to the discovery of different motivational patterns than considering each factor separately.

Motivational distance was the only construct in the current study which was significantly predicted by engagement in ISLPs while itself not being a significant predictor of engagement. Specifically, affective, cognitive, and linguistic engagement significantly predicted changes in the participants' motivational distance over the course of the study. Affective and linguistic engagement were associated with significant *increases* in the motivational distance, meaning that participants felt further removed from their ideal future selves over time. This suggests that engaging in highly enjoyable informal activities, as well as focusing explicitly on some of the language features within them, might make learners more aware of the gaps in their own language knowledge. This awareness may cause learners both to feel less confident in their current abilities (current L2 self) and to adjust their expectations of what it would mean to be a successful language user (future L2 self), thus leading to an overall increase in the motivational distance.

Cognitive engagement, by contrast, was associated with a *decrease* in the distance between the current and ideal future L2 selves. A possible explanation can be found, once again, in the qualitative data. A number of focus group participants who did not engage extensively in ISLPs explained that they did not enjoy doing certain activities in English (especially reading books or watching serials) because they did not understand enough of the meaning to feel fully immersed in the activity.²⁴

²⁴ The participants' descriptions of this are reminiscent of the concept of *flow*, which is describes a state in which learners experience optimal levels of challenge, interest, and focus (e.g. Dewaele & MacIntyre, 2019; Egbert, 2003).

Conversely, then, this implies that feeling more immersed in an informal activity (higher cognitive engagement) might be interpreted by some pupils as a sign of their own language competence, thus decreasing the distance between the current and ideal future L2 self.

10.3.3 Attitudes towards Informal and Formal Learning

Of all the motivational constructs included in this study, the participants' attitudes towards informal and formal language learning were most strongly related to engagement in ISLL. On the one hand, this is evident in the finding that these factors were the only variables to share significant relationships with the overall engagement patterns (LPA groups), even after the inclusion of proficiency in the same model. In addition, both variables shared significant direct relationships with multiple aspects of engagement.

Attitudes towards informal learning

Looking first at the individual engagement dimensions, the effects of the attitudes towards informal learning were largely as expected: First, learners with stronger positive beliefs about ISLPs and their effectiveness as language learning tools were more interested in and gained more enjoyment from engaging in informal activities (affective engagement). Second, informal attitudes also predicted greater behavioural engagement both in terms of the quantity and, to a lesser extent, the variety of ISLPs. This is in accordance with the results of Lai, Hu, and Lyu's (2018) study, who used SEM to show that the extent to which university-level EFL learners in Hong Kong and the USA engaged in entertainment- and information-oriented activities (roughly

equivalent to ISLPs in this study) was directly predicted by their perceptions of the usefulness of these practices as tools for language learning.

Positive attitudes towards informal learning were also associated with higher levels of cognitive engagement (how focused and absorbed pupils felt while engaging in an informal activity), but there was no significant relationship between informal learning beliefs and linguistic engagement. This is in contrast to Cole's (2015) theory about the role of learner perceptions in ISLL, based on his observation that the 'fully autonomous self-directed learners' in his sample were more likely than the 'classroom-trained learners' to try to 'figure out' the meaning of new language features when they engaged in informal activities (p. 286). Connecting this observation with his insights into the two groups' perceptions of ISLL, Cole proposed that learners' attitudes towards ISLPs (specifically acknowledging their potential to promote language learning, instead of viewing them primarily as leisure activities) would predict the mental effort that they would invest into decoding the language in informal activities. The current study's findings, however, do not support Cole's theory, given that attitudes towards informal learning were not significantly associated with linguistic engagement. This is after no evidence was found in support of Cole's second hypothesis, that greater mental effort (cognitive/linguistic engagement) was associated with larger gains in language proficiency (Section 10.2.2.5).

Given that attitudes towards informal learning had significant effects on behavioural and affective engagement, as discussed above, it is not surprising that they also emerged as significant predictors of the participants' overall patterns engagement in ISLPs, as represented by the LPA groups. Specifically, participants who believed more strongly in the effectiveness of ISLPs were more likely to be part of the

first group, who engaged with a broader range of informal activities, spent more time on them, were more emotionally engaged and focused less on the linguistic features. Remarkably, this effect was approximately equal in strength to that of prior proficiency and together these two variables explained 32% of the variance in the participants' general engagement patterns.

Attitudes towards formal learning

Attitudes towards formal learning at T1 had significant effects on affective and cognitive engagement that were similar in direction and size to the effects of informal attitudes. However, unlike attitudes towards informal learning, formal learning beliefs also positively predicted linguistic engagement.

The higher levels of cognitive and linguistic engagement associated with positive beliefs about formal language learning could be an indication that these participants were trying to apply learning strategies they had been taught to use in school in their informal second language practices. This interpretation is accordance with previous research on language learning beyond the classroom, which has suggested that learners who strongly believe in the efficacy of formal learning may transfer the structured approaches they were taught in the classroom to their informal practices (Chik, 2014, 2018; Lai, 2015, see also Section 10.3.1. above).

Meanwhile, the greater sense of interest in and enjoyment of informal activities (affective engagement) experienced by participants with strong positive beliefs in formal learning might indicate that these students felt positively surprised to find ways of applying their language knowledge outside of the classroom. The same reason was suggested earlier, in an effort to explain the positive effect of instrumental obligation on affective engagement (Section 10.3.1). A slightly different interpretation

is provided by Chik (2018), who reported that learners who, in their out-of-class language practices, deliberately chose resources and employed learning strategies to try to re-create more structured classroom learning experiences also felt a sense of ‘guilty pleasure’ (p. 54) when engaging in more unstructured activities, such as watching TV shows or film in their second language. However, it seems more likely that the higher levels of affective engagement indicate that the ‘threshold of interest’ that an informal L2 activity had to cross in the participants’ minds to be worth engaging in was higher for participants with more positive attitudes towards formal learning (cf. *instrumental obligation*, Section 10.3.1).

Lastly, although none of the individual engagement dimensions significantly affected the learners’ attitudes over the course of this study, the general engagement patterns (LPA groups) significantly predicted changes in the participants’ attitudes towards formal learning. This effect was small and negative, implying that the students in the first LPA group (who engaged in significantly more, and more varied, ISLPs, and were more affectively but less linguistically engaged) had become less positive in their attitudes towards formal learning and the types of language skills taught in the classroom.²⁵

This might be interpreted as evidence for what researchers have called the widening ‘authenticity gap’ (Henry, 2013) between the methods for teaching and learning widely used in classrooms and the ways in which individuals use and improve their language skills in their free time, but could also be indicative of students developing a more internal locus of control/becoming more autonomous in their learning. This is supported by finding that the attitudes towards formal and informal

²⁵ Or, conversely, that the pupils in Group 2 believed more strongly in formal instruction by the end of this study.

learning affected engagement in ISLPs in separate and unique ways, which shows that a preference for one mode of learning does not necessarily preclude learners from making an effort in the other context. As argued previously in this thesis (Section 10.2.1), instead of contrasting formal and informal learning in terms of their relative effectiveness, viewing both sets of practices as complimentary may be of benefit to language learners and teachers, as well as to the directions for future research in the field of informal second language learning.

11. Conclusion

This last chapter begins by providing a summary of the aims and key findings of this study. The discussion then turns towards the strengths and limitations of the study, followed by directions for future research and implications for language learning and teaching.

11.1 Key Findings

Patterns of engagement in informal second language learning

This study has drawn upon the concept of *engagement* as a way to systematically study the nature of informal second language practices, not merely in terms of the behaviour, but also of the thoughts and feelings of learners who participate in them. This involved surveying the informal activities of a large sample of year 10 students at *Gymnasium*-type secondary schools in Germany, who were advanced learners of English as a second language.

The informal second language practices of the sample in this study, in terms of types of activities and time spent on them (behavioural engagement), were largely similar to those reported in previous studies of ISLL with different populations. Most of the students spent around two hours per day on informal practices, with listening to music, browsing social media, and watching online videos and serials being the most popular activities. How interesting and enjoyable an activity seemed to them was the main criterion by which learners appeared to choose whether or not to engage in an informal activity. This explains why the recorded levels of affective engagement were generally high. How much effort the students expected they would need to invest (cognitive engagement) in order to reach a good level of comprehension was also an important criterion. In general, the learners paid little

attention to form during informal practices (linguistic engagement), but some would occasionally choose to focus on particular language features if it was essential for their overall understanding. Furthermore, it was discovered that reading (except on social media) and activities which involved language production were associated with higher levels of cognitive and linguistic engagement, supporting a distinction (proposed by Sundqvist & Sylvén, 2016) between these activities as more ‘active’ as compared to ‘passive activities’ such as listening to music or watching serials and films.

Engagement and proficiency

The primary aim in exploring different aspects of engagement in informal second language practices was to understand their links with language proficiency, language learning motivation, and attitudes. The study also attempted to investigate the directionality of these relationships, that is to explore the extent to which more proficient and motivated students were more likely to engage in informal second language practices, and the extent to which these practices promoted language development and changes in motivation.

Thus, one of the key findings of the current study was that engagement in informal second language practices did not significantly predict changes in general language proficiency over the course of the study. The learners’ prior proficiency did, however, predict a significant proportion of the variance in the participants’ overall engagement patterns, as well as in the individual dimensions of behavioural, cognitive, and linguistic engagement when these were considered separately. Again, it is important to contextualise this finding: While the methods used in this

study do not directly provide evidence for the directional relationship between these variables, the temporal order of the data supports the interpretation that the effect of proficiency on informal second language practices is stronger than in the opposite direction. However, given some of the limitations of the current study (see Section 11.3), the hypothesis that ISLPs can promote second language acquisition should not be rejected without evidence from further studies (Section 11.4).

Prior proficiency had different effects on cognitive and linguistic engagement as compared to behavioural engagement. In particular, higher-proficiency learners tended to participate in more diverse English-language activities and spent greater amounts of time using the language outside of school and in addition to any formally assigned tasks. However, they also tended to dedicate less cognitive effort to informal activities and focus less on the language features they encountered.

Thus, setting aside the issue of the directionality of the effects discovered in the current study (as discussed above), this study found a positive link between the quantity and variety of ISLPs and proficiency, but a negative link between proficiency and cognitive and linguistic engagement. This provides evidence against the hypothesis that investing greater mental effort into ISLPs, and particularly focusing on decoding and learning specific language features, is associated with greater language proficiency. The findings therefore imply that any development in language proficiency derived from informal second language practices that were to be found in future studies are likely be mostly incidental.

The relationships between proficiency and engagement were somewhat mediated by the motivational variables, meaning that the direct effects of proficiency weakened

after also accounting for motivation. This was strongest in the case of the cognitive and linguistic engagement: The direct negative effect of proficiency on these two dimensions of engagement was only revealed in the last model, which included both proficiency and motivation variables. This is because concurrently modelling the effects of proficiency and motivation allowed for the separation of the direct negative and indirect positive effects (via motivation) of proficiency on engagement, which had previously cancelled out each other in the model that only included proficiency and engagement.

Motivation and engagement

Altogether, this study has provided some evidence for a reciprocal relationship between motivation and engagement in informal second language practices. That is, all of the engagement variables were directly affected by some aspects of the students' motivation or attitudes towards language learning. In turn, engagement in informal activities also predicted changes in some of these variables.

Attitudes towards formal and informal language learning related most strongly to engagement in ISLPs. Positive beliefs about informal learning in particular had a strong effect on the overall engagement patterns. This effect was comparable in strength to that of prior language proficiency. This variable also individually affected all dimensions of engagement except for linguistic engagement. This constitutes a particularly interesting finding as it shows that learners who believed more strongly that informal activities promoted language learning did not invest more effort into language processing (contrary to what was hypothesised by Cole, 2015).

Another surprising finding was that positive attitudes towards formal learning, as well as a strong instrumental obligation (feeling externally pressured to learn the L2), predicted higher levels of affective, cognitive, and linguistic engagement in ISLPs. The link to affective engagement could mean that students who generally rely more on formal language instruction (display an 'external locus of control' in language learning, cf. Benson, 2011) derive greater enjoyment from informal activities as a sort of 'guilty pleasure' (as suggested by Chik, 2018). However, it seems more likely that this finding results from these students being more hesitant to engage in ISLPs generally, choosing only to do so when they were highly interested in a specific activity. Meanwhile, the higher levels of cognitive and linguistic engagement which were associated with a strong instrumental obligation and positive attitudes towards formal learning could suggest that these pupils were attempting to transfer some of the structured approaches that were familiar to them from the classroom to their informal practices (as observed previously by Chik, 2014, 2018; Lai, 2015).

In return, engagement in informal second language practices also predicted changes in attitudes towards formal learning over time. More specifically, students who tended to be behaviourally and affectively more engaged, but paid less attention to language, became less positive in their attitudes towards formal learning over the course of the study. This provides some tentative evidence for the hypothesis that frequent informal language contact can de-motivate students in formal language instruction (create an 'authenticity gap', according to Henry, 2013; Lamb & Arisandy, 2019).

In addition to instrumental orientation (discussed above), two other language learning orientations had significant positive effects on various aspects of engagement in informal second language practices: knowledge orientation (viewing English as relevant to the pursuit of personal interests and new information) and intrinsic motivation (an inherent affinity for the language and the learning process). Generally speaking, this is in accordance with findings from previous studies of motivation in ISLL (Cole, 2015; Lyrigkou, 2016). Meanwhile, instrumental benefits and international orientation were not related in any way to informal language practices in this study.

The pupils' perceptions of their present language abilities (current L2 self) did not share a significant link with informal second language practices when the learners' actual proficiency was also taken into account. This finding stands in contrast to previous studies, which have reported positive correlations between informal language contact and self-efficacy or confidence in language learning, albeit without also taking into account actual language abilities (Lai et al., 2015; J. S. Lee, 2019; Sundqvist, 2009, 2011). The future L2 self, however, shared a weak reciprocal relationship with one engagement measure, even after accounting for proficiency: Pupils with a strong desire to become competent future L2 users were more likely to use a greater variety of informal resources; in turn, using English in a variety of ways beyond school also helped to strengthen the learners' desire for language competence. A similar link between the ideal future self and informal practices was also reported in a previous ISLL study with adult learners in Indonesia (Lamb & Arisandy, 2019).

Finally, the strongest effects of engagement in ISLPs in this study were on motivational distance (how far removed participants felt from their ideal future self). Higher levels of affective and linguistic engagement were associated with an *increase* in this distance, suggesting that enjoying informal activities and focusing on the language within them can make learners more aware of gaps in their language knowledge. Conversely, stronger cognitive engagement predicted a decrease in motivational distance, suggesting that learners gained confidence in their own language abilities if they felt more immersed in an informal activity and had to make less of an effort to understand the general meaning.

11.2 Strengths of the Study

This section briefly reviews the two primary methodological strengths of the current study: the well-targeted nature of the data collection instruments, which were developed specifically for the informal second language learning context, and the comprehensiveness and sophistication of the quantitative analytic approach.

Because the questionnaires used to measure engagement, motivation, and attitudes available in previous research often made explicit reference to features of classroom-based learning, designing new instruments that were more suitable for use in the informal second language learning context was an essential part of this study. This involved a bottom-up, mixed-methods approach, wherein qualitative data were collected first, which subsequently informed the design of the quantitative instruments. The interviews in Phase I allowed the researcher to explore what the learners themselves thought were the most relevant aspects of their engagement in informal practices; their motivation for engaging in these activities and for language

learning more generally; as well as their beliefs about the effectiveness of ISLL as compared to formal, school-based learning. These findings facilitated the adaptation of appropriate items from existing motivation, attitudes, and engagement questionnaires, wherever possible, as well as the addition of new items where necessary. The qualitative data were particularly valuable in showing the kind of language that students from the target population used to talk about their experiences, which in turn helped to ensure comprehensible phrasing (or re-phrasing) of the questionnaire items.

The motivation and attitudes questionnaire, informal language learning diary, and C-tests were all piloted with another group of pupils drawn from the same target population and were refined based on the analysis of these data in Phase II. The refined instruments were then used to collect the large-scale quantitative data set in Phase III. Preliminary analyses were conducted on this data set to confirm the measurement validity of the instruments, before moving on to the main statistical analyses. This comprehensive approach meant that, altogether, each instrument was rigorously tested through Exploratory Factor Analysis, Confirmatory Factor Analysis, and finally Longitudinal Measurement Invariance Modelling (C-test and motivation and attitudes questionnaire only, which were administered at two time points).

The sophistication of the main statistical analyses equally added to the quality of this study. Modelling all of the psychological constructs (engagement, motivation, attitudes, proficiency) as latent variables made it possible to statistically separate the measurement error inherent in self-reported data. Using Structural Equation Modelling (SEM) also allowed the researcher to account for the fact that the participating students were not sampled at random, but clustered in classrooms,

effectively controlling for effects of teacher/classroom practices and group dynamics. Similarly, using Multi-Level SEM to analyse the engagement data enabled the separation of the inter-individual from the intra-individual variation (differences between participants vs. variation between different diary entries by the same participant), which allowed the researcher to derive more accurate measures of each participant's general levels of engagement in ISLPs. In addition, this approach treats intra-individual differences as substantively interesting data rather than as error variance, which led to some illuminating insights about the different functions and interactions of the various engagement aspects at the activity- and participant levels.

Finally, employing SEM methods instead of more traditional approaches (e.g. regression analyses) made it possible to test more complex models involving mediation and multiple outcome variables. Thus, this study was able to investigate a number of direct and indirect links simultaneously (e.g. direct effects of proficiency on engagement and indirect effects via motivation and attitudes), while also exploring the possible directions of these relationships.

11.3 Limitations

In addition to the methodological strengths, this study has a number of limitations in its design and analytical approach. Since many have already been noted at the appropriate points throughout this thesis, this section contains only a brief discussion of the main limitations of this study as a whole, as well as of its qualitative and quantitative components.

A limitation which relates to the collection of both the qualitative and quantitative data for this study is the reliance on self-report measures, which have a number of

known weaknesses (see e.g. Hoskin, 2012). For example, participants may have differed in how they interpreted certain interview questions or survey items, or in their understanding and use of the cognitive differential rating scales (e.g. *Completely agree* — *Don't agree at all*). There may also have been instances of social desirability bias, particularly in the focus group interviews. That is, some pupils may have felt reluctant to talk about engaging in activities which they were afraid might be perceived negatively by their peers (e.g. as 'nerdy' or 'uncool') or by the researcher. Although these are certainly issues to be aware of, it is difficult to envisage a way of measuring variables such as engagement, motivation, and attitudes that does not involve asking the learners themselves. In addition, as mentioned above (Section 11.2), this study used latent variable modelling to statistically account for some of the measurement error inherent in the quantitative self-report data.

The biggest limitation which was unique to the qualitative part of this study was related to the participant sample. Due to the time-intensive nature of the focus group interviews, only one school (two classes) could be recruited to participate. In addition, the sample had a marked gender imbalance, with girls making up almost 80% of the participants. On the one hand, this was because the participating school had an 'uncommonly high' proportion of girls in their year 10 classes, according to the headmaster. On the other hand, more of the boys declined to take part in the study than the girls (possible reasons were discussed in Section 4.1). These sampling limitations should be kept in mind when considering the qualitative results, especially relating to gender and the nature of formal learning experiences. However, as shown throughout the discussion (Chapter 10), many of these findings could be triangulated

using the quantitative data, which stemmed from a larger, more gender-balanced sample of students at nine different schools.

Several limitations of the quantitative part of this study related to the nature of the data collection instruments. The first concerns the use of an event-contingent diary to collect data on engagement in informal second language practices (see Section 6.1.3). This method was chosen (over one-off, retrospective self-report questionnaires, which have been used in most previous ISLL studies) as the researcher thought that it would provide a more accurate record of the pupils' experiences while engaging in ISLPs, especially their concurrent cognitive and emotional states. Yet, it cannot be ascertained how representative the activities recorded over the period of the study were of the participants' general ISLPs. This is an important consideration, given that both previous research (particularly the longitudinal case-studies conducted by Kussyk, 2017b) and the qualitative findings in this study suggest that informal practices can vary substantially from day to day as well as across longer periods of time.

Against this background, another limitation is that not all pupils who participated in the informal language learning diary submitted entries across seven days, as requested. This was presumably because this data collection method placed a high demand on the participants in terms of time and self-directed effort at home. To investigate the extent to which the study's findings were influenced by participants who submitted diary entries across fewer than five days, several additional tests were conducted (see Sections 9.1.3.5 and 9.3.3). These consistently supported the conclusion that the statistical results were not substantially affected by the number of days across which diary entries were recorded. Nevertheless, in the cases where

limited data were available, there is also less certainty that the recorded data were indeed representative of pupils' general informal practices.

Another limitation in the quantitative data collection relates to the use of C-tests to measure language proficiency (see also Section 10.2.1.1). As rather general measures, these tests may not have been sensitive enough to fully capture the changes in the participants' language abilities over the course of the study. Relatedly, the period between the pre- and post-tests (14 weeks) may not have been long enough for the participants to make significant progress in their general language development. Both of these limitations constitute possible explanations for the lack of significant predictors of changes in language proficiency over time in this study.

The C-tests were developed specifically for use with the sampled population of language learners, following the recommendation of one of the leading experts on C-testing (Grotjahn, 2002). It must be acknowledged that this (as well as the decision to develop new motivation and attitudes questionnaires) inevitably involved a trade-off against the established reliability of existing, more widely validated instruments. Care was taken to extensively pilot the new instruments and to demonstrate that the data were sufficiently valid for the subsequent analyses (see Section 11.2). Nevertheless, it must be acknowledged that, through these rigorous tests, some systematic differences were revealed between the proficiency scores of two sub-samples: the pupils who took test Version 1 first vs. those who first completed Version 2. Due to the split-sample, cross-over design it could not be determined whether this was because the two test versions were not completely equal in difficulty, or whether the participant sub-samples differed in terms of their actual

proficiency. If the tests did differ in difficulty, this may have contributed to the lack of substantial measurable changes in language proficiency over the course of this study.

Lastly, the size of the data set which formed the basis for the analyses in Phase III must be mentioned. The sophisticated statistical methods used here (see Section 11.2) required a larger sample size than more traditional regression models due to the higher number of parameters that needed to be estimated. While the researcher successfully recruited a very large sample of 354 students, only 213 (60%) completed the proficiency test and motivation questionnaire at both time points, as well as the informal language learning diary (see Section 9.1.1). By using maximum likelihood modelling to impute missing values wherever possible and adopting a two-step Factor Score Path Modelling approach, it was possible to carry out the planned analyses while meeting the commonly accepted minimum sample-size requirements (see Section 9.3.2). Nevertheless, as statisticians have observed that error rates can be inflated when conducting SEM analyses with smaller sample sizes (e.g. Bentler & Chou, 1987; R. B. Kline, 2015; Wolf et al., 2013), a more complete data set would have been desirable in the current study.

In future research, higher sample sizes might be obtained by reducing the burden on the individual participants. For example, depending on the constructs which are to be measured, it may be possible to further limit the number of questions included in the diary, therefore decreasing the time it takes for participants to complete. Additional measures may also be taken to further reduce the self-directed effort involved in completing the informal language learning diary. Although it would require extensive effort on the part of the researcher, creating an app for data collection, instead of directing participants to an online questionnaire, may make the

instrument feel less cumbersome for the participants to use. Such a dedicated tool could also considerably simplify the process of sending survey reminders. The daily notifications, which the researcher sent via WhatsApp, seemed to substantially improve participation rates in Phase III of this study (see Section 8.2). Therefore, it is possible that slightly more frequent reminders may prove to be even more effective in getting participants involved.

11.4 Directions for Future Research

The findings and the limitations of the current study highlighted several directions for future research, which would be of particular value to the field of informal second language learning at this point. In particular, there are at least three categories of future studies which could directly expand upon the work presented here: (1) comparative studies of ISLL, involving more than one learner population; (2) longitudinal and cross-sectional studies; and (3) studies of the internal psychological processes at work during ISLL. Each of these ideas will be briefly introduced below.

First and foremost, it must be mentioned that the scope of this study was necessarily limited: targeted sampling and exclusions from the analyses were employed in order to control, in as much as possible, for the influence of variables such as first language background, second language experience, formal instruction, and socioeconomic background (see Section 3.3). However, it seems likely that the informal second language practices and language-related perceptions of any sample would be affected, in part, by socio-contextual variables. For example, since English-language popular culture, as well as the language itself, played an important role in these participants'

lives, the findings may not be readily transferable to contexts in which English is less dominant as an L2. The results of this study therefore require further validation through future research involving different learner populations and importantly, target languages other than English.

Although attempts were made throughout the discussion to compare the results of this study with findings from previous research, it is unclear to what extent any differences observed may be affected by variations in population characteristics, or differences in the methodological approaches used. In order to be able to draw valid conclusions about the effects of variables such as gender, age, level of current language proficiency (i.e. higher vs. lower-proficiency learners), and educational, national, or cultural context, large-scale comparative studies are needed, which employ unified measures of engagement in and outcomes of informal second language practices.

In addition, the field could also benefit from more cross-sectional and longitudinal studies. The former could focus on capturing general trends in informal language use, similar to the annual UK language trends survey (Tinsley, 2019), which provides an annual snapshot of formal learning and teaching, thus allowing changes over time to be observed and highlighted. Meanwhile, longitudinal studies could be used to track developments in informal second language practices within the same group of learners, and any associated changes in variables such as proficiency and motivation, in the same population over time (cf. Olsson & Sylvén, 2015). Despite the lack of such findings in the current study, future research may well show that the various dimensions of engagement can predict developments in language proficiency, and thus explain some of the individual and group differences reported in the literature,

for example with regard to gender (Kuppens, 2007, 2010; Sundqvist, 2009, 2011; Sundqvist & Sylvén, 2014) and current level of proficiency (higher vs. lower level learners Lai et al., 2018). Such studies may also explore the possibility that engagement in informal second language practices and language abilities might share a reciprocal, rather than unidirectional relationship, that is, exerting a positive influence over one another over time.

Finally, building on the insights which this study has provided into the ways in which individual learners engage in ISLPs, researchers may wish to increase their focus on the internal psychological processes that influence language learning moment-to-moment, including situated motivation and conscious attention to language. At several points throughout this thesis, the situated and dynamic nature of engagement was discussed. For example, several focus group participants described how their attention could be captured suddenly by a salient language feature, or that they might choose, spontaneously, to employ particular strategies to increase their understanding, and in turn their enjoyment, of the activity in which they were currently engaged. In other words, the different aspects of engagement, and their relationship with other variables, seemed to vary not only from learner to learner, but also across time and according to features of the informal activities themselves. In the quantitative analyses, multi-level modelling helped to account for this variability and allowed the researcher to more accurately study the inter-individual differences in informal engagement. Understanding patterns at this broad level may serve as a starting point for understanding the processes involved in informal second language learning.

Thus, future studies should also investigate intra-individual variations in engagement: how learners' emotional involvement, attention distribution ('noticing of language features') and strategic thinking may fluctuate between different informal practices or even within the same activity, and how such variations relate to characteristics of the activities and learners. Researchers may also wish to explore how far such micro-level processes can help to explain individual differences in long-term outcomes of informal second language contact.

Finally, there are likely to be variables other than those that were investigated here, which directly or indirectly influence informal practices and learning outcomes. These may include factors such as the perceived attitudes of teachers, peers and parents; previous formal and informal learning experiences; time spent abroad; individual learning preferences, strategies, and needs, etc. It is clearly beyond the scope of the present project to cover all of these aspects. With informal second language learning being a relatively young field of study, there are plenty of other questions left to explore in future research.

11.5 Implications for Teaching and Learning

This study has provided further statistical evidence of the significant positive link between informal second language practices and proficiency which has been widely reported in the ISLL literature (Cole, 2015; De Wilde et al., 2019; Jurkovič, 2019; Kuppens, 2007, 2010; Lyrigkou, 2016; Olsson, 2011; Sundqvist, 2009, 2011; M. H. Verspoor et al., 2011), although it has also called into question the directionality of this relationship. Meanwhile, the qualitative data suggested that informal activities could contribute to second language abilities in ways which complement what is

learned in school: Whereas the interviewed learners widely attributed their foundational language knowledge, and in particular their command of grammar and syntax, to formal language instruction, many pupils also thought that ISLPs had helped them expand and individualise their vocabularies; given them time to improve and strengthen some of their skills (primarily reading and listening comprehension) and an opportunity to familiarise themselves with more authentic language, as opposed to the formal register taught in school. The vast amounts of time that students have been shown to spend on informal practices in this and other studies must also be considered, especially compared to the few hours per week that are dedicated language learning and teaching in schools.

Finally, as indicated in various parts of this study as well as prior research (Cole, 2015; Lai et al., 2015; Lamb & Arisandy, 2019; J. S. Lee, 2019; Sundqvist, 2009, 2011), there seems to be a positive association between students' engagement in informal practices and various aspects of language learning motivation, including self-confidence, a desire to learn, and concrete learning goals. All of these findings suggest that it may be valuable to view informal and formal approaches to language learning as complementary and that promoting informal second language contact would be a worthwhile endeavour for researchers, teachers, and policy makers.

While the current study's findings may not prove to be universally applicable, they add to our general understanding of the characteristics of informal second language practices themselves and of those students who are more likely to engage in them. This, in turn, can provide an idea of the types of informal second language engagement we may wish to promote, and a starting point for thinking about how best to

encourage other learners to take part in informal activities who are not yet as involved as some of their peers.

Proficiency was found to be most strongly linked with the time that students spent on informal activities and the number of different practices in which they engaged. The importance of engaging in a variety of activities has also been highlighted in previous research (Lai et al., 2015; J. S. Lee, 2017, 2019; J. S. Lee & Dressman, 2017). This suggests that promoting ISLPs should involve not only encouraging learners to spend more time with their second language beyond the classroom, but also helping them to diversify their repertoire of informal experiences.

The qualitative data suggested that in this particular student population, engagement in informal activities involving language production primarily depended on the pupils' personal circumstances, that is, whether they had familial relations or friendships with other users of English. Thus, helping learners to foster interpersonal relationships with other speakers of their target language may be an effective way to encourage productive language use outside of the classroom. This could be aided through more formal approaches, such as tandem learning programs (Cziko, 2004), or more informally through helping learners take advantage of online interest groups and other such communities (Cole, 2015; Sockett, 2014b).

Many students who did not engage in receptive activities, on the other hand, seemed to experience difficulties with listening comprehension, and a smaller number mentioned finding it difficult to read longer texts in English. Indeed, the quantitative data also revealed that proficiency significantly predicted the students' likelihood of engaging in listening activities (including listening to music, watching films and TV, listening to spoken English, and watching videos on social media), as well as the time spent on reading (books, online texts, as well as social media content).

Thus, teachers wanting to promote more informal language contact among their students may wish to focus on improving their students' confidence in their own listening and reading abilities and/or spend more time on improving these skills.

In addition to the positive association with behavioural engagement, there was a negative link between proficiency and cognitive and linguistic engagement in the current study. Although the directionality of this relationship still needs further confirmation, it could suggest that it may be more beneficial for students to seek out activities in which they can immerse themselves easily and not to focus extensively on consciously trying to learn new language features. In addition, the qualitative data suggested that comprehension difficulties and a prominent focus on form decreased learners' interest in and enjoyment of informal activities and negatively impacted upon their desire to engage further in ISLPs. These findings highlight the importance of choosing activities at an appropriate level of difficulty, and which cater to the learners' specific interests.

Lai (2015) found that explicitly discussing and encouraging out-of-class learning seemed to have little impact on students' actual practices but, much like Sundqvist and Sylvén (2016), suggests providing learners with specific suggestions for resources and guidance on how to use them. Similarly, in Phase I of the current study, some students (who seemed to be less confident in their language abilities than their peers) commented that they wished for their teachers or school librarians to provide them with more suggestions for materials (e.g. films, books, serials) with which they could engage outside of school, or better yet, to provide them with more guidance on how to find and select materials appropriate for their current level of proficiency. This need for specific guidance is something that teacher trainers and

policy makers should also be aware of, as it suggests that instructors need to be trained to recognise, utilise, and communicate the potential positive effects of informal second language practices for teaching and learning.

Sundqvist and Sylvén (2016) proposed a *Thirty Day Extramural English Challenge* as a way of introducing learners to new ways of using their second language outside of the classroom. The researchers compiled a list of 44 informal practices (some broader, like ‘watch[ing] a movie in English with no subtitles’, some linked to specific resources such as ‘practic[ing] lyrics from a [favourite] song ... at lyricstraining.com’) and suggested that teachers ask their students to engage in, and reflect upon, one activity each day for one month. This method seems to hold the potential for broadening learners’ informal repertoires while also catering to a variety of different interests and fostering autonomy by letting students decide which activities to engage with each day.

In order to help students take charge of their own learning even more, teachers may also wish to involve the learners themselves in the process of designing the list of resources. This will provide an additional opportunity to discuss the potential benefits of different types of informal practices while also ensuring that the suggestions will feel relevant to the specific group of learners, which is likely to increase the likelihood that the activity will inspire learners to continue to use their second language outside of school after completing the Thirty Day Challenge.

As concluded by Ushioda (2011), the goal should be to ‘engage learner identities and interests [during] lessons and promote a sense of continuity between what they learn and do in the classroom, and who they are and what they are interested in doing in their lives outside the classroom’ (p. 204). This should also help to address the

emergence of a potential ‘authenticity gap’ (Henry, 2013; Henry & Cliffordson, 2017; Lamb & Arisandy, 2019), for which there was some evidence in this study, wherein active informal second language users may feel demotivated by the perceived mismatch between the language and activities which they experienced in the classroom and beyond.

Finally, modelling a positive attitude towards informal learning could be an important step in increasing students’ second language contact outside of the classroom. After all, beliefs about the effectiveness of informal learning were found to be an important predictor (equal in strength to proficiency) of overall patterns of engagement in ISLPs in this study. In addition to this, the results of the current study suggest that perceptions about the target language itself have great effects on engagement in ISLPs, meaning that targeting these beliefs inside the classroom may allow teachers to exert some positive influence over learners’ behaviour outside of school.

In particular, it is recommended that teachers specifically try to influence the views of students who perceive their second language mainly as a school subject, something which has to be taught, and which they are obligated to learn. This study suggested that there may be great benefits to promoting learner autonomy (decreasing instrumental orientation and strengthening the future L2 self), showing students that using their second language can be enjoyable (increasing intrinsic motivation) and can empower them to pursue their personal interests, broaden their horizons, and develop new ideas (strengthening knowledge orientation).

In summary, the findings presented and discussed here suggest that there are strong relationships between informal activities and L2 proficiency, motivation, and

attitudes which researchers, teachers, and policy makers need to be aware of in order to be able to take advantage of the (still largely untapped) potential of informal practices for SLA research and pedagogy.

References

- Ainley, M., & Patrick, L. (2006). Measuring self-regulated learning processes through tracking patterns of student interaction with achievement activities. *Educational Psychology Review*, 18, 267–286.
- Akbari, E., Ahmad, N., Simons, R.-J., & Pilot, A. (2016). Student engagement and foreign language learning through online social networks. *Asian-Pacific Journal of Second and Foreign Language Education*, 1(4), 1–22.
- Al-Hoorie, A. H. (2017). Sixty years of language motivation research: Looking back and looking forward. *SAGE Open*.
- Albrecht, J., Fink, P., & Tiemann, H. (2015). *Ungleiches Deutschland: Sozioökonomischer Disparitätenbericht 2015*. Bonn.
- Alderson, J. C. (2002). Testing proficiency and achievement: principles and practice. In J. A. Coleman, R. Grotjahn, & U. Raatz (Eds.), *University language testing and the C-test* (pp. 15–30). Bochum: AKS-Verlag.
- Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of School Psychology*, 44, 427–445.
- Ariani, M. G., & Ghafournia, N. (2016). The relationship between socio-economic status, general language learning outcome, and beliefs about language learning. *International Education Studies*, 9(2), 90–98.
- Arras, U., Eckes, T., & Grotjahn, R. (2002). C-Tests im Rahmen des “Tests Deutsch als Fremdsprache” (TestDaF): Erste Forschungsergebnisse. In R. Grotjahn (Ed.), *Der C-Test. Theoretische Grundlagen und praktische Anwendungen. Bd. 4.* (pp. 175–209). Bochum: AKS-Verlag.
- Asoodar, M., Atai, M. R., Vaezi, S., & Marandi, S. S. (2014). Examining effectiveness of communities of practice in online English for academic purposes (EAP) assessment in virtual classes. *Computers & Education*, 70, 291–300.
- Babaii, E., & Ansary, H. (2001). The C-test: A valid operationalization of the reduced redundancy principle? *System*, 29, 209–219.
- Bailly, S. (2011). Teenagers learning languages out of school: What, why and how do they learn? How can school help them? In P. Benson & H. Reinders (Eds.), *Beyond the language classroom* (pp. 119–131). Basingstoke: Palgrave MacMillan.
- Bandalos, D. L. (2014). Relative performance of categorical diagonally weighted least squares and robust maximum likelihood estimation. *Structural Equation Modeling*, 21(1), 102–116.
- Bauer, D. J., & Curran, P. J. (2004). The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. *Psychological Methods*, 9, 3–29.
- Benson, P. (2011a). Language learning and teaching beyond the classroom: An introduction to the field. In P. Benson & H. Reinders (Eds.), *Beyond the language classroom* (pp. 7–16). Basingstoke: Palgrave MacMillan.
- Benson, P. (2011b). *Teaching and researching autonomy in language learning* (2nd ed.). London: Pearson Education.
- Benson, P., & Chik, A. (2010). New literacies and autonomy in foreign language learning. In M. J. Luzón, M. N. Ruiz-Madrid, & M. L. Villanueva (Eds.), *Digital genres, new literacies, and autonomy in language learning* (pp. 63–80). Cambridge Scholars Publishing.
- Benson, P., & Reinders, H. (2011). *Beyond the language classroom*. Basingstoke: Palgrave MacMillan.
- Bentler, P. M., & Chou, C.-P. (1987). Practical issues in Structural Modeling. *Sociological Methods & Research*, 16, 78–117.
- Bialosiewicz, S., Murphy, K., & Berry, T. (2013). *An introduction to measurement invariance testing: Resource packet for participants*. Washington, DC.
- Black, R. W. (2005). Access and affiliation: The literacy and composition practices of English-language learners in an online fanfiction community. *Journal of Adolescent & Adult Literacy*, 49, 118–128.
- Black, R. W. (2006). Language, culture, and identity in online fanfiction. *E-Learning*, 3, 170–184.
- Black, R. W. (2009a). English-language learners, fan communities, and 21st-century skills. *Journal of Adolescent & Adult Literacy*, 52(8), 688–697.

- Black, R. W. (2009b). Online fan fiction, global identities, and imagination. *Research in the Teaching of English*, 43(4), 397–425.
- Blunch, N. J. (2013). *Introduction to structural equation modeling using IBM SPSS Statistics and Amos* (2nd ed.). London: Sage Publications.
- Boekaerts, M. (2016). Engagement as an inherent aspect of the learning process. *Learning and Instruction*, 43, 76–83.
- Bogards, S., & Raatz, U. (2002). Sind C-Tests trainierbar? In R. Grotjahn (Ed.), *Der C-Test. Theoretische Grundlagen und praktische Anwendungen. Bd. 4*. Bochum: AKS-Verlag.
- Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, 54, 579–616.
- Bonke, J. (2003). Paid work and unpaid work: Diary information versus questionnaire information. *Social Indicators Research*, 70(3), 349–368.
- Boo, Z., Dörnyei, Z., & Ryan, S. (2015). L2 motivation research 2005–2014: Understanding a publication surge and a changing landscape. *System*, 55, 145–157.
- Böwing, C., Frome, W., Gerking, K., Haupt, D., & Schulte, W. (2006). *Kerncurriculum für das Gymnasium Schuljahrgänge 5–10*. Hannover.
- Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, 36, 111–150.
- Byrne, B. M. (2010). *Structural equation modeling with AMOS* (2nd ed.). New York: Routledge.
- Cadierno, T., & Eskildsen, S. W. (2015). *Usage-based perspectives on second language learning*. Berlin: De Gruyter.
- Cambridge English Qualifications: A2 Flyers. (2019). Retrieved August 14, 2019, from <https://www.cambridgeenglish.org/exams-and-tests/flyers/>
- Carlson Wurpts, I. (2012). *Testing the limits of Latent Class Analysis*. Arizona State University.
- Celeux, G., & Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification*, 13, 195–212.
- Chamot, A. U. (2004). Issues in language learning strategy research and teaching. *Electronic Journal of Foreign Language Teaching*, 1(1), 14–26.
- Chen, J. F., Warden, C. A., & Cheng, H. (2005). Motivators that do not motivate: The case of Chinese EFL learners and the influence of culture on motivation. *TESOL Quarterly*, 39(4), 609–633.
- Chik, A. (2014). Digital gaming and language learning: Autonomy and community. *Language Learning and Technology*, 18(2), 85–100.
- Chik, A. (2018). Learning a language for free: Space and autonomy in adult foreign language learning. In G. Murray & T. E. Lamb (Eds.), *Space, place, and autonomy in language learning* (pp. 56–72). London: Routledge.
- Christenson, S. L., Reschly, A., & Wylie, C. (Eds.). (2012). *Handbook of research on student engagement*. New York: Springer.
- Clément, R., & Kruidenier, B. G. (1983). Orientations in second language acquisition: I. The effects of ethnicity, milieu and target language on their emergence. *Language Learning*, 33(3), 273–291.
- Cohen, A. D. (1998). *Strategies in learning and using a second language* (Longman). Harlow.
- Cole, J. (2015). *Foreign language learning in the age of the internet*. Unpublished doctoral thesis. University of Oxford.
- Cole, J., & Vanderplank, R. (2016). Comparing autonomous and class-based learners in Brazil: Evidence for the present-day advantages of informal, out-of-class learning. *System*, 61, 31–42.
- Coleman, J. A. (2002). The European proficiency survey: An overview of findings. In J. A. Coleman, R. Grotjahn, & U. Raatz (Eds.), *University language testing and the C-test* (pp. 201–206). Bochum: AKS-Verlag.
- Conrath, C., Higgins, C. A., & McClean, R. J. (1983). A comparison of the reliability of questionnaire versus diary data. *Social Networks*, 5(3), 315–322.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge. Retrieved from <https://rm.coe.int/1680459f97>
- Cox, C. B. (2014). *21st century skills and principles of flor in the foreign language classroom*. Master's thesis. Brigham Young University.

- Csizér, K., & Kormos, J. (2009). Learning experiences, selves, and motivated behaviour: a comparative analysis of structural models for Hungarian secondary and university learners of English. In *Motivation, language identity and the L2 self* (pp. 98–119). Bristol: Multilingual Matters.
- Csizér, K., & Kormos, J. (2014). The ideal L2 self, self-regulatory strategies and autonomous learning: A comparison of different groups of English language learners. In K. Csizér & M. Magid (Eds.), *The impact of self-concept on language learning* (pp. 73–88). Bristol: Multilingual Matters.
- Cziko, G. A. (2004). Electronic tandem language learning (eTandem): A third approach to second language learning for the 21st century. *CALICO Journal*, 22(1), 25–39.
- Daller, H., & Phelan, D. (2006). *The C-test and TOEIC as measures of students' progress in intensive short courses in EFL*. (R. Grotjahn, Ed.). Bern: Peter Lang.
- De Bot, K., & Larsen-Freeman, D. (2011). Researching second language development from a Dynamic Systems Theory perspective. In Marjolijn H. Verspoor, K. De Bot, & W. Lowie (Eds.), *A dynamic approach to second language development* (pp. 5–23). Amsterdam: John Benjamins.
- de Freitas, S. (2013). *Learning in immersive worlds: A review of game-based learning*. Bristol.
- De Wilde, V., Brysbaert, M., & Eyckmans, J. (2019). Learning English through out-of-school exposure. Which levels of language proficiency are attained and which types of input are important? *Bilingualism: Language and Cognition*, 1–15.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behaviour*. New York: Plenum.
- Devlinger, I., & Rosseel, Y. (2017). Factor score path analysis: An alternative for SEM? *Methodology*, 13, 31–38.
- Dewaele, J.-M., & MacIntyre, P. D. (2019). Flow in the Spanish foreign language classroom. *Preprint*. Retrieved from https://www.researchgate.net/publication/331546936_Flow_in_the_Spanish_Foreign_Language_Classroom
- Die niedersächsischen allgemein bildenden Schulen in Zahlen*. (2016). Hannover. Retrieved from https://www.mk.niedersachsen.de/startseite/service/statistik/allgemein_bildende_schulen/die-niedersaechsischen-allgemein-bildenden-schulen-in-zahlen-6505.html
- DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation*, 14(20), 1–11.
- Dörnyei, Z. (1994). Motivation and motivating in the foreign language classroom. *The Modern Language Journal*, 78(3), 273–284.
- Dörnyei, Z. (2001). *Motivational strategies in the language classroom*. Cambridge: Cambridge University Press.
- Dörnyei, Z. (2003). Attitudes, orientations, and motivations in language learning: Advances in theory, research, and applications. In Z. Dörnyei (Ed.), *Attitudes, orientations, and motivations in language learning* (pp. 3–32). Oxford: Blackwell.
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Erlbaum.
- Dörnyei, Z. (2009a). *Questionnaires in second language research* (2nd ed.). London: Routledge.
- Dörnyei, Z. (2009b). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Bristol: Multilingual Matters.
- Dörnyei, Z. (2015). Future self-guides and vision. In K. Csizér & M. Magid (Eds.), *The impact of self-concept on language learning* (pp. 7–18). Bristol: Multilingual Matters.
- Dörnyei, Z., & Ottó, I. (1998). Motivation in action: A process model of L2 motivation. *Working Papers in Applied Linguistics*, 4, 43–69.
- Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. New York: Routledge.
- Dörnyei, Z., & Taguchi, T. (2009). Illustration: Developing a motivation questionnaire. In Z. Dörnyei (Ed.), *Questionnaires in second language research* (2nd ed., pp. 111–126). London: Routledge.

- Dörnyei, Z., & Ushioda, E. (2011). *Teaching and Researching Motivation*. Harlow: Pearson Education.
- Dunn, L. M., & Dunn, D. M. (2007). PPVT-4: Peabody picture vocabulary test. Minneapolis: Pearson Assessments.
- Eccles, J. S. (2016). Engagement: Where to next? *Learning and Instruction*, 43, 71–75.
- Eccles, J. S., & Wang, M.-T. (2012). So what is student engagement anyway? In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 133–145). New York: Springer.
- Eckes, T., & Grotjahn, R. (2006). A closer look at the construct validity of C-tests. *Language Testing*, 23(3), 290–325.
- Egbert, J. (2003). A study of flow theory in the foreign language classroom. *The Modern Language Journal*, 87(4), 499–518.
- Eisinga, R., te Grotenhuis, M., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health*, 58(4), 637–642.
- Ellis, N. C. (2017). Cognition, corpora, and computing: Triangulating research in usage-based language learning. *Language Learning*, 67(1), 40–65.
- Elswood, K. (2015). Popular culture, the culture of the 21st century. Retrieved May 21, 2019, from <https://www.shoutoutuk.org/2015/09/11/meaning-western-culture/>
- Ethical guidelines for educational research. (2011). Retrieved from <http://www.bera.ac.uk/wp-content/uploads/2013/12/BERA-Ethical-Guidelines-2011.pdf>
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272–299.
- Field, A. P. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). London: Sage Publications.
- Filmer, D., & Pritchett, L. (1999). The effect of household wealth on educational attainment: Evidence from 35 countries. *Population and Development Review*, 25, 85–120.
- Finn, J. D., & Zimmer, K. S. (2012). Student engagement: What is it? Why does it matter? In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 97–132). New York: Springer.
- Finney, S. J., & DiStefano, C. (2008). Non-normal and categorical data in structural equation modeling. In G. R. Hancock & R. D. Mueller (Eds.), *Structural Equation Modeling: A second course* (pp. 269–314). Charlotte, NC: Information Age Publishing.
- Flowerday, T., & Schraw, G. (2003). Effect of choice on cognitive and affective engagement. *The Journal of Educational Research*, 96(4), 207–215.
- Foorman, B. R., Petscher, Y., Stanley, C., & Truckenmiller, A. (2017). Latent profiles of reading and language and their association with standardized reading outcomes in Kindergarten through tenth grade. *Journal of Research on Educational Effectiveness*, 10(3), 619–645.
- Forero, C. G., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2009). Factor analysis with ordinal indicators: A Monte Carlo study comparing DWLS and ULS estimation. *Structural Equation Modeling*, 16(4), 625–641.
- Forsman, L. (2004). *Language, culture and context: Exploring knowledge and attitudes among Finland-Swedish EFL-students with particular focus on extracurricular influence*. Master's thesis. Åbo Akademi, Vasa.
- Fredericks, J. A., Blumenfeld, P. C., Friedel, J., & Paris, A. H. (2005). School engagement. In K. A. Moore & L. Lippman (Eds.), *What do children need to flourish? Conceptualizing and measuring indicators of positive development*. New York: Springer Science and Business Media.
- Fredericks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74, 59–109.
- Gardner, R. C. (1985a). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.
- Gardner, R. C. (1985b). *The Attitude/Motivation Test Battery: Technical report*. London, CA.
- Gardner, R. C. (2000). Correlation, causation, motivation and second language acquisition. *Canadian Psychology*, 41(1), 10–24.
- Gardner, R. C. (2010). *Motivation and second language acquisition: The Socio-Educational Model*. New York: Peter Lang.

- Gardner, R. C., Tremblay, P. F., & Masgoret, A.-M. (1997). Toward a full model of second language learning: An empirical investigation. *The Modern Language Journal*, 81, 344–362.
- Geoghegan, L. (2018). International posture, motivation and identity in study abroad. In C. Pérez Vidal, S. López-Serrano, J. Ament, & D. J. Thomas-Wilhem (Eds.), *Learning context effects: Study abroad, formal instruction and international immersion classrooms* (pp. 215–253). Berlin: language science press.
- Godfroid, A., Housen, A., & Boers, F. (2010). A procedure for testing the Noticing Hypothesis in the context of vocabulary acquisition. In M. Pütz & L. Scola (Eds.), *Inside the learner's mind: Cognitive processing and second language acquisition* (pp. 169–197). Amsterdam: John Benjamins.
- Godfroid, A., & Winke, P. (2015). Investigating implicit and explicit processing using L2 learners' eye-movement data. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 325–348). Amsterdam: John Benjamins.
- Godwin-Jones, R. (2019). Riding the digital wilds: Learner autonomy and informal language learning. *Language Learning & Technology*, 23(1), 8–25.
- Green, C. F. (1999). Categorising motivational drives in second language acquisition. *Language, Culture and Curriculum*, 12(3), 265–270.
- Greenhow, C., Robelia, B., & Hughes, J. E. (2009). Learning, teaching, and scholarship in the digital age: Web 2.0 and classroom research: What path should we take now? *Educational Researcher*, 38, 246–259.
- Grenfell, M., & Macaro, E. (2007). Claims and critiques. In A. D. Cohen & E. Macaro (Eds.), *Language learner strategies: Thirty years of research and practice* (pp. 9–28). Oxford: Oxford University Press.
- Grotjahn, R. (2002). Konstruktion und Einsatz von C-Tests: Ein Leitfaden für die Praxis. In R. Grotjahn (Ed.), *Der C-Test. Theoretische Grundlagen und praktische Anwendungen. Bd. 4.* (pp. 211–225). Bochum: AKS-Verlag.
- Hannibal Jensen, S. (2019). Language learning in the wild: A young user perspective. *Language Learning & Technology*, 23(1), 72–86.
- Hastings, A. J. (2002). In defence of C-testing. In R. Grotjahn (Ed.), *Der C-Test. Theoretische Grundlagen und praktische Anwendungen. Bd. 4.* (pp. 11–25). Bochum: AKS-Verlag.
- Hedges, L. V., & Rhoads, C. (2010). *Statistical power analysis in education research*. Washington, DC.
- Henry, A. (2013). Digital games and ELT: Bridging the authenticity gap. In E. Ushioda (Ed.), *International perspectives on motivation* (pp. 135–155). Basingstoke: Palgrave MacMillan.
- Henry, A., & Cliffordson, C. (2017). The impact of out-of-school factors on motivation to learn English: Self-discrepancies, beliefs, and experiences of self-authenticity. *Applied Linguistics*, 38(5), 713–736.
- Hessel, G. (2011). *The legacy of the year abroad: The motivational development of modern languages finalists at an English university from a "possible L2 selves" perspective*. Master's thesis. University of Oxford.
- Hessel, G. (2015). From vision to action: Inquiring into the conditions for motivational capacity of ideal second language selves. *System*, 52, 103–114.
- Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, 94, 319–340.
- Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, 1, 104–121.
- Ho, D. G. E. (2006). The focus group interview: Rising to the challenge of qualitative research methodology. *Australian Review of Applied Linguistics*, 29(1), 05.01–05.19.
- Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and a meta-analysis. *Sociological Methods & Research*, 26(3), 329–367.
- Horowitz, E. K. (1988). The beliefs about language learning of beginning university foreign language students. *Modern Language Journal*, 72, 283–294.
- Hoskin, R. (2012). The dangers of self-report. Retrieved January 17, 2017, from <http://www.sciencebrainwaves.com/the-dangers-of-self-report/>
- Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.

- Huitt, W. (2011). Motivation to learn: An overview. In *Educational psychology interactive*. Valdosta State University.
- Hull, G., & Schultz, K. (2001). Literacy and learning out of school: A review of theory and research. *Review of Educational Research*, 71, 575–611.
- Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258–286). Cambridge: Cambridge University Press.
- Internet-based research (IBR): Best practice guidance 06_Version 3.0*. (2016). Oxford. Retrieved from http://www.admin.ox.ac.uk/media/global/wwwadminoxacuk/localsites/curec/documents/BPG_06_Internet-Based_Research.pdf
- Iwaniec, J. (2014). Motivation of pupils from southern Poland to learn English. *System*, 45, 67–78.
- Janosz, M. (2012). Part IV commentary: Outcomes of engagement and engagement as outcome: Some consensus, divergences, and unanswered questions. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 695–703). New York: Springer.
- Jhangiani, R., Tarry, H., & Stangor, C. (2014). Exploring Attitudes. In *Principles of Social Psychology – 1st International Edition*. Victoria, BC: BCCampus.
- Jurkovič, V. (2019). Online informal learning of English through smartphones in Slovenia. *System*, 80, 27–37.
- Kaiser, H. F. (1970). A second-generation little jiffy. *Psychometrika*, 30, 401–415.
- Kaiser, H. F., & Rice, J. (1974). Little jiffy, mark IV. *Educational and Psychological Measurement*, 34(1), 120–138.
- Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research*, 44(3), 486–507.
- Khodadady, E. (2014). Construct validity of C-tests: A factorial approach. *Journal of Language Teaching and Research*, 5(6), 1353–1362.
- Kim, T.-Y., & Kim, Y.-K. (2014). EFL students' L2 motivational self system and self-regulation: Focusing on elementary and junior high school students in Korea. In K. Csizér & M. Magid (Eds.), *The impact of self-concept on language learning* (pp. 87–107). Bristol: Multilingual Matters.
- Kinginger, C. (2009). *Language learning and study abroad: A critical reading of research*. London: Palgrave MacMillan.
- Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). London and New York: Routledge.
- Kline, R. B. (2015). *Principles and practice of structural equation modeling*. New York: Guildford Press.
- Knobel, M., & Lankshear, C. (2007). *A new literacies sampler*. New York: Peter Lang.
- König, C. (2018). Bildungspläne/Lehrpläne der Bundesländer für allgemeinbildende Schulen.
- Kormos, J., Kiddle, T., & Csizér, K. (2011). Systems of goals, attitudes, and self-related beliefs in second-language-learning motivation. *Applied Linguistics*, 35(2), 495–516.
- Krueger, R. A., & Casey, M. A. (2015). *Focus groups: A practical guide for applied research* (5th ed.). Thousand Oaks, CA: Sage Publications.
- Kuppens, A. H. (2007). De invloed van mediagebruik op de verwerving van Engelse woordenschat: een empirische studie bij Vlaame jongeren. *Tijdschrift Voor Communicatiewetenschap*, 35(4), 325–336.
- Kuppens, A. H. (2010). Incidental foreign language acquisition from media exposure. *Learning, Media and Technology*, 35(1), 65–85.
- Kusyk, M. (2017a). *Les dynamiques du développement de l'anglais au travers d'activités informelles en ligne: une étude exploratoire auprès d'étudiants français et allemands*. Doctoral thesis. Université de Strasbourg & Pädagogische Hochschule Karlsruhe.
- Kusyk, M. (2017b). The development of complexity, accuracy, and fluency in L2 written production through informal participation in online activities. *CALICO Journal*, 34(1), 75–96.

- Kusyk, M., & Sockett, G. (2012). From informal resource usage to incidental language acquisition: Language uptake from online television viewing in English. *La Revue Du Groupe d'Étude et de Recherche En Anglais de Spécialité*, 62, 45–65.
- Lai, C. (2015). Perceiving and traversing in-class and out-of-class learning: Accounts from foreign language learners in Hong Kong. *Innovation in Language Learning and Teaching*, 9(3), 265–284.
- Lai, C., Hu, X., & Lyu, B. (2018). Understanding the nature of learners' out-of-class language learning experience with technology. *Computer Assisted Language Learning*, 31(1–2), 114–143.
- Lai, C., Zhu, W., & Gong, G. (2015). Understanding the quality of out-of-class English learning. *TESOL Quarterly*, 49(2), 278–308.
- Lam, W. S. E. (2000). L2 literacy and the design of the self: A case study of a teenager writing on the internet. *TESOL Quarterly*, 34(3), 457–482.
- Lam, W. S. E. (2004). Second language socialization in bilingual chat room: Global and local considerations. *Language Learning & Technology*, 8(1), 44–65.
- Lam, W. S. E. (2006). Re-envisioning language, literacy, and the immigrant subject in new mediascapes. *Pedagogies: An International Journal*, 1, 171–195.
- Lamb, M. (2004). Integrative motivation in a globalizing world. *System*, 32(1), 3–19.
- Lamb, M., & Arisandy, F. E. (2019). The impact of online use of English on motivation to learn. *Computer Assisted Language Learning*.
- Lambert, C., Philp, J., & Nakamura, S. (2016). Learner-generated content and engagement in second language task performance. *Language Teaching Research*, 1–16.
- Laufer, B., & Nation, I. S. P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16, 33–51.
- Lee-Ellis, S. (2009). The development and validation of a Korean C-Test using Rasch analysis. *Language Testing*, 26(2), 245–274.
- Lee, H.-G. (2012). *ESL learners' motivation and task engagement in technology enhanced language learning contexts*. Doctoral thesis. Washington State University.
- Lee, J. S. (2017). Informal digital learning of English and second language vocabulary outcomes: Can quantity conquer quality? *British Journal of Education Technology*.
- Lee, J. S. (2019). Quantity and diversity of informal digital learning of English. *Language Learning & Technology*, 23(1), 114–126.
- Lee, J. S., & Dressman, M. (2017). When IDLE hands make an English workshop: Informal digital learning of English and language proficiency. *TESOL Quarterly*, 52(2), 435–445.
- Lee, J. S., & Lee, K. (2019). Informal digital learning of English and English as an international language: The path less traveled. *British Journal of Education Technology*, 50, 1447–1461.
- Leow, R. P., & Zamora, C. C. (2017). Intentional and incidental L2 learning. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition*. London: Routledge.
- Leverett, T. (2012). Motivation and Second Language Acquisition: The Socio-Educational Model (Review). *Teaching English as a Second or Foreign Language*, 16(2).
- Levine, G. S. (2003). Student and instructor beliefs and attitudes about target language use, first language use, and anxiety: Report of a questionnaire study. *Modern Language Journal*, 87, 343–364.
- Lightbown, P. M., & Spada, N. (2013). Views about language learning. In *How languages are learned* (4th ed., pp. 3–4). Oxford: Oxford University Press.
- Litt, M. D., Cooney, N. L., & Morse, P. (1998). Ecological momentary assessment (EMA) with treated alcoholics: methodological problems and potential solutions. *Health Psychology*, 17, 48–52.
- Liu, C.-C., Wang, P.-C., & Tai, S.-J. D. (2016). An analysis of student engagement patterns in language learning facilitated by Web 2.0 technologies. *ReCALL*, 28(2), 104–122.
- Liu, X. (2016). Cross-lagged relations between motivation and proficiency in English as a foreign language among Chinese university students. *Journal of Education and Practice*, 7(6), 29–37.
- Loewen, S., Li, S., Fei, F., Thompson, A., Nakatsukawa, K., Ahn, S., & Chen, X. (2009). Second language learners' beliefs about grammar instruction and error correction. *Modern Language Journal*, 93, 91–104.

- Lyriqkou, C. (2016). *Young learners' engagement with English in their free time through informal sources. How do frequency of exposure, effort and motivation affect oral production?* Master's thesis. University of Oxford.
- Lyriqkou, C. (2018). Not to be overlooked: agency in informal language contact. *Innovation in Language Learning and Teaching*, 1–16.
- Macaro, E. (2006). Strategies for language learning and for language use: Revising the theoretical framework. *The Modern Language Journal*, 90(3), 320–337.
- MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modification in covariance structure analysis: The problem of capitalisation on chance. *Psychological Bulletin*, 111, 450–504.
- MacIntyre, P. D., Mackinnon, S. P., & Clément, R. (2009a). The baby, the bathwater, and the future of language learning motivation research. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 43–65). Bristol: Multilingual Matters.
- MacIntyre, P. D., Mackinnon, S. P., & Clément, R. (2009b). Toward the development of a scale to assess possible selves as a source of language learning motivation. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 193–214). Bristol: Multilingual Matters.
- Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, 41, 954–969.
- Marsh, H. W., Lüdtke, U., & Trautwein, J. S. A. (2009). Classical latent profile analysis of academic self-concept dimensions: Synergy of person- and variable-centered approaches to theoretical models of self-concept. *Structural Equation Modeling*, 16, 191–225.
- Masgoret, A.-M., & Gardner, R. C. (2003). Attitudes, motivation, and second language learning: A meta-analysis of studies conducted by Gardner and associates. *Language Learning*, 53(1), 123–163.
- McAlister, A. M., Lee, D. M., Ehler, K. M., Kajfez, R. L., Faber, C. J., & Kennedy, M. S. (2017). *Qualitative coding: An approach to assess inter-rater reliability* (No. 18189). Washington, DC.
- Mercer, S. (2011). *Towards an understanding of language-learner self-concept*. London: Springer.
- Mills, N. (2011). Situated learning through social networking communities: The development of join enterprise, mutual engagement, and a shared repertoire. *CALICO Journal*, 28(2), 345–368.
- Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345–368). Newbury Park, CA: Sage.
- Muthén, L. K., & Muthén, B. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, 4, 599–620.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. (C. A. Chapelle & S. Hunston, Eds.). Cambridge: Cambridge University Press.
- Noels, K. A. (2001a). Learning Spanish as a second language: Learners' orientations and perceptions of their teachers' communication style. *Language Learning*, 5(1), 107–144.
- Noels, K. A. (2001b). New orientations in language learning motivation: Towards a model of intrinsic, extrinsic, and integrative orientations and motivation. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 43–68). Honolulu: University of Hawai'i.
- Noels, K. A., Pelletier, L. G., Clément, R., & Vallerand, R. J. (2000). Why are you learning a second language? Motivational orientations and self-determination theory. *Language Learning*, 53(1), 57–85.
- Nunan, D., & Richards, J. C. (2015). *Language learning beyond the classroom*. New York: Routledge.
- Olsson, E. (2011). *"Everything I read on the Internet is in English" – On the impact of extramural English on Swedish 16-year-old pupils' writing proficiency*. Doctoral thesis. University of Gothenburg.
- Olsson, E., & Sylén, L. K. (2015). Extramural English and academic vocabulary: A longitudinal study of CLIL and non-CLIL students in Sweden. *Journal of Applied Language Studies*, 9(2), 77–103.

- Oyserman, D., & James, L. (2009). Possible selves: From content to process. In K. Markman, W. M. P. Klein, & A. Suhr (Eds.), *The handbook of imagination and mental stimulation* (pp. 373–394). New York: Psychology Press.
- Parrish, M. (2016). Is YouTube replacing traditional television? Retrieved May 18, 2019, from <https://www.theodysseyonline.com/youtube-is-replacing-traditional-television>
- Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104, 1–15.
- Price, P. C., Jhangiani, R. S., & Chiang, I. A. (2015). *Research methods in Psychology* (2nd ed.). Vancouver: BCcampus Open Ed.
- Privacy Shield List. (2018). Washington, DC. Retrieved from <https://www.privacyshield.gov/list>
- Qualtrics Research Suite. (2018). Retrieved from <http://www.qualtrics.com>
- Raatz, U., & Klein-Braley, C. (1981). The C-test – A modification of the cloze procedure. In T. Culhane, C. Klein-Braley, & D. K. Stevenson (Eds.), *Practice and problems in language testing*. Colchester: University of Essex, Department of Language and Linguistics Occasional Papers No. 26.
- Raatz, U., & Klein-Braley, C. (2002). Introduction to language testing and to C-tests. In J. A. Coleman, R. Grotjahn, & U. Raatz (Eds.), *University language testing and the C-test* (pp. 75–91). Bochum: AKS-Verlag.
- Ramaswamy, V., DeSarbo, W. S., Reibstein, D. J., & Robinson, W. T. (1993). An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Science*, 12(1), 103–124.
- Reichert, M., Brunner, M., & Martin, R. (2014). Do test takers with different language backgrounds take the same C-test? The effect of native language on the validity of C-tests. In R. Grotjahn (Ed.), *Der C-Test. Aktuelle Tendenzen* (pp. 109–136). Frankfurt (Main): Peter Lang.
- Renninger, K. A., & Bachrach, J. E. (2015). Studying triggers for interest and engagement using observational methods. *Educational Psychologist*, 50(1), 58–69.
- Reschly, A. L., & Christenson, S. L. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 3–20). New York: Springer.
- Rosenberg, J. M. (2018). Introduction to tidyLPA. Retrieved November 20, 2018, from https://jrosen48.github.io/tidyLPA/articles/Introduction_to_tidyLPA.html#software-approach-to-carrying-out-lpa-interface-to-mclust-and-to-mplus
- Rosseel, Y. (2018). Multilevel SEM model syntax. Retrieved November 16, 2018, from <http://lavaan.ugent.be/tutorial/multilevel.html>
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54–67.
- Ryan, S. (2009). Self and identity in L2 motivation in Japan: The ideal L2 self and Japanese learners of English. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 120–143). Bristol: Multilingual Matters.
- Salomon, G. (1984). Television is “easy” and print is “tough”: The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology*, 76(4), 647–658.
- Sarapuu, I., & Alas, E. (2016). Developing a c-test to measure language ability as an alternative to a skills-based test. *Estonian Papers in Applied Linguistics*, 12, 237–252.
- Sauro, S., & Zourou, K. (2019). What are the digital wilds? *Language Learning & Technology*, 23(1), 1–7.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–158. Retrieved from <http://applij.oxfordjournals.org/content/11/2/129.abstract>
- Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, 29(4), 304–321.
- Seidlhofer, B. (2005). Key concepts: English as a lingua franca. *ELT Journal*, 59(4), 339–341.
- Shiffman, S., Hufford, M., Hickcox, M., Paty, J. A., Gnys, M., & Kassel, J. D. (1997). Remember that? A comparison of real-time versus retrospective recall of smoking lapses. *Journal of Consulting and Clinical Psychology*, 65, 292–300.

- Sigott, G. (2004). *Towards identifying the C-test construct*. Bern: Peter Lang.
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75(3), 417–453.
- Skinner, E. A., Kindermann, T. A., Connell, J. P., & Wellborn, J. G. (2009). Engagement as an organizational construct in the dynamics of motivational development. In K. Wentzel & A. Widfield (Eds.), *Handbook of motivation at school* (pp. 223–245). Mahwah, NJ: Erlbaum.
- Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualisation and assessment of children's behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement*, 69(3), 493–525.
- Sockett, G. (2014a). Language acquisition from online informal learning: Analysis of written production. In *The online informal learning of English* (pp. 102–112). London: Palgrave MacMillan.
- Sockett, G. (2014b). *The online informal learning of English*. London: Palgrave MacMillan.
- Sockett, G. (2014c). The practices of OILE: Listening to MOD. In *The online informal learning of English* (pp. 39–51). London: Palgrave MacMillan.
- Sockett, G. (2014d). The practices of OILE: Quantitative surveys of the activities associated with OILE. In *The online informal learning of English* (pp. 31–39). London: Palgrave MacMillan.
- Sockett, G., & Toffoli, D. (2012). Beyond learner autonomy: a dynamic systems view of the informal learning of English in virtual online communities. *ReCALL*, 24(2), 138–151.
- Stevick, E. W. (1990). Research on what? Some terminology. *Modern Language Journal*, 74, 143–153.
- Sugita McEown, M., Noels, K. A., & Everhart Chaffee, K. (2014). At the interface of the Socio-Educational Model, Self-Determination Theory and the L2 Motivaitonal Self System Models. In K. Csízér & M. Magid (Eds.), *The impact of self-concept on language learning* (pp. 19–50). Bristol: Multilingual Matters.
- Sundqvist, P. (2009). *Extramural English matters: Out-of-school English and its impact on Swedish ninth graders' oral proficiency and vocabulary*. Doctoral thesis. Karlstad University.
- Sundqvist, P. (2011). A possible path to progress: Out-of-school English language learners in Sweden. In P. Benson & H. Reinders (Eds.) (pp. 106–118). Basingstoke: Palgrave MacMillan.
- Sundqvist, P., & Olin-Schneller, C. (2013). Classroom vs. extramural English: Teachers dealing with demotivation. *Language and Linguistics Compass*, 7(6), 329–338.
- Sundqvist, P., & Sylvé, L. K. (2014). Language-related computer use: Focus on young L2 English learners in Sweden. *ReCall*, 26(1), 3–20.
- Sundqvist, P., & Sylvé, L. K. (2016). *Extramural English in teaching and learning*. London: Palgrave MacMillan.
- Sundqvist, P., & Wikström, P. (2015). Out-of-school digital gameplay and in-school L2 English vocabulary outcomes. *System*, 51, 65–76.
- Sykes, J. M., Oksoz, A., & Thorne, S. L. (2008). Web 2.0, synthetic immersive environments, and mobile resources for language education. *CALICO Journal*, 25(3), 528–546.
- Tabachnik, B. G., & Fidell, L. S. (2014). *Using multivariate statistics* (6th ed.). Harlow: Pearson Education.
- Taboada Barber, A., Gallagher, M., Smith, P., Buehl, M. M., & Beck, J. S. (2016). Examining student cognitive and affective engagement and reading instructional activities: Spanish-speaking English learners' reading profiles. *Literacy Research and Instruction*, 55(3), 209–236.
- Thorne, S. L., Black, R. W., & Sykes, J. M. (2009). Second language use, socialization, and learning in internet interest communities and online gaming. *The Modern Language Journal*, 93, 802–821.
- Thurston, M. (2015). International posture, attitudes and motivation among mainland Chinese EFL learners in Singapore. *Malaysian Journal of ELT Research*, 11(2), 1–16.
- Tinsley, T. (2019). *Language Trends*. British Council. London. Retrieved from <https://www.britishcouncil.org/research-policy-insight/research-reports/language-trends-2019>
- Tódor, E., & Dégi, Z. (2016). Language attitudes, language learning experiences and individual strategies: What does school offer and what does it lack? *Acta Universitatis Sapientiae, Philologica*, 8(2), 123–137.

- Toffoli, D., & Sockett, G. (2010). How non-specialist students of English practice informal learning using web 2.0 tools. *La Revue Du Groupe d'Étude et de Recherche En Anglais de Spécialité*, 58, 125–154. Retrieved from <https://asp.revues.org/1851>
- Toni, A., & Rostami, M. (2012). Overall motivation and the promotion of EFL learners' oral proficiency. *Theory and Practice in Language Studies*, 2(11), 2336–2341.
- Turek, R. (2016). What YouTube looks like in a day. Retrieved May 21, 2019, from <https://medium.com/@synopsi/what-youtube-looks-like-in-a-day-infographic-d23f8156e599>
- Usage of content languages for websites. (2019).
- Ushioda, E. (2009). A person-in-context relational view of emergent motivation, self and identity. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 215–228). Bristol: Multilingual Matters.
- Ushioda, E. (2011). Language learning motivation, self, and identity: Current theoretical perspectives. *Computer Assisted Language Learning*, 24(3), 199–210.
- van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, 9(4), 486–492.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4–70.
- Vandergrift, L. (2005). Relationships among motivation orientations, metacognitive awareness and proficiency in L2 listening. *Applied Linguistics*, 26(1), 70–89.
- Verspoor, M., De Bot, K., & van Rein, E. (2011). English as a foreign language: The role of out-of-school language input. In A. de Houwer & A. Wilton (Eds.), *English in Europe today: Sociocultural and educational perspectives* (pp. 147–166). Amsterdam: John Benjamins.
- Verspoor, M. H., de Bot, K., & Lowie, W. (Eds.). (2011). *A dynamic approach to second language development: methods and techniques*. Methods. Amsterdam: John Benjamins.
- Wang, M.-T., Fredricks, J. A., Ye, F., Hofkens, T. L., & Schall Lin, J. (2016). The math and science engagement scales: development, validation, and psychometric properties. *Learning and Instruction*, 43, 16–26.
- Wheeler, L., & Reis, H. T. (1997). Self-recording of everyday life events: Origins, types, and uses. *Journal of Personality*, 59, 339–354.
- White, C. J. (1995). Autonomy and strategy use in distance foreign language learning: Research findings. *System*, 23(2), 297–221.
- Williams, C. (2016). Traditional TV viewing is over: YouTube habit is permanent, warns researchers. Retrieved January 18, 2016, from <https://www.telegraph.co.uk/finance/newsbysector/mediatechnologyandtelecoms/media/12067340/Traditional-TV-viewing-is-over-YouTube-habit-is-permanent-warn-researchers.html>
- Williams, G. A., & Kibowski, F. (2015). Latent class analysis and latent profile analysis. In L. A. Jason & D. S. Glenwick (Eds.), *Handbook of methodological approaches to community-based research: Qualitative, quantitative, and mixed-methods* (pp. 143–152). Oxford: Oxford University Press.
- Wo stehen die Bundesländer beim Bildungscheck? [How do the federal states rank on the educational assessment?]. (2016). Retrieved May 5, 2017, from <http://www.insm-bildungsmonitor.de>
- Wolf, E. J., Harrington, K. M., Clark, Shaunna, L., & Miller, M. W. (2013). Sample size requirements for Structural Equation Models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, 76(6), 913–934.
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practice. *The Counseling Psychologist*, 34(6), 806–838.
- Wunsch, G., Russo, F., & Mouchart, M. (2010). Do we necessarily need longitudinal data to infer causal relationships. *Bulletin de Méthodologie Sociologique*, 106, 5–18.
- Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *Modern Language Journal*, 86, 54–66.

- Yashima, T. (2009). International posture and the ideal L2 self in the Japanese EFL context. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 144–163). Bristol: Multilingual Matters.
- Zhang, L. J., & Xiao, Y. (2006). Language learning strategies, motivation, and EFL proficiency: A study of Chinese tertiary-level non-English majors. *Asian Englishes*, 9(2), 20–47.
- Zunke, M., Althoff, H. K., Apel, J., Lässig, H. S., Schültke, L., & Schredl, M. (2015). Comparing questionnaire and diary measures for eliciting nightmare frequency. *International Journal of Dream Research*, 8(2), 129–134.

Appendix A. Methods Phase I

A.1 Focus Group Interview Guide

A.1.1 English

Greeting and General Guidelines

Thank you for agreeing to take part in this study and for taking the time for this focus group. As you will have heard from your teacher, my aim in this study is to find out more about how students like you use English in their everyday lives, outside of school. Hopefully, the results of my study will help to improve language learning for future students like you.

I would like to remind you that I will audio-record our conversation using a Dictaphone and my smartphone, but all data are confidential and will be kept safe and anonymous.

I have a few other guidelines for you:

1. If there are any questions that you don't understand or that you would like me to repeat, please let me know.
2. Although I'm interested in knowing as much detail as possible, I do not wish to pry into your personal life, so feel free to skip any questions that you do not feel comfortable answering.
3. There are no right or wrong answers. I expect that all of you will have different experiences and opinions. Please feel free to tell us about your experiences and opinions even if — and especially if — they are different from somebody else's.

Finally, I don't want you to feel like you always have to speak to me: you can answer to what somebody else has said, agree with them, or disagree. I am only here to ask you questions, to listen, and to make sure that everybody has a chance to participate in the conversation.

If at any time you feel that you want to withdraw from this interview, you're completely free to do so. In that case, I will delete any record of our conversation up to that point.

Do you have any questions for me at this point? Alright, let's begin.

Questions

1. First, please tell me a little about how you use English in your free time (besides in school or when you're doing homework).
 - a. [For example, do you ever read in English, listen to music, watch films or TV series, or use English on social networking sites like Facebook or Twitter?]¹
 - b. Do you sometimes use websites or apps to learn English — such as Duolingo, Babbel, or Memrise — that your teacher didn't tell you to use?
 - c. [Is there anything else?]
2. What do you like about these activities?
 - a. [This doesn't have to be about the language. For example...]
3. Could you estimate how much time per week you spend on these activities?
 - a. Is this more or less the same every week, or does it vary?
 - b. If yes, why?
4. So from what I'm hearing, I get the feeling that you mostly *read/listen to*² English. Can you think of occasions where you communicate by writing or speaking English?
 - a. [For example, do you ever write blog posts, Facebook status updates, or tweet in English, or do you write comments on YouTube or Facebook?]
 - b. [Do you talk to people in English on Facebook or Skype?]
5. Why do you *read/watch TV or films/surf the Internet/etc.* in English rather than German?
6. Please tell me about how good your understanding of English in these circumstances is. This is not a test! I am interested in hearing about how you evaluate your own language skills. In this way, I can get a more complete picture of how and why you use English in your free time.
 - a. How important is it to you that you understand everything when you *watch TV/read/play games/etc.* in English?
 - i. Why/why not?
 - b. Is there anything you'd like to do *watch/read/etc.*, but you don't because you feel that you don't understand enough English?

¹ Square brackets indicate questions or comments that can be added in, to prompt participants to talk more extensively on a particular issue.

² Italics indicate parts of questions that should be adapted to the individual group, based on what they have previously said. For example, if a group of participants talked mainly about watching TV series in English, the subsequent questions would reference mostly this activity.

7. Do you ever consciously focus on specific language features – e.g. words, phrases, or grammatical constructions – while you are *watching television/listening to music/reading/etc.*?
 - a. Why/why not?
 - b. How often would you say this happens?
8. What do you do when you notice a particular word, phrase, or construction?
 - a. Is there anything specific that you do to learn the English words or grammar that you encounter in this way, or to remember them better?
9. Does this depend on the device you're using? For example, whether you are on the computer versus on your phone, or when you are *reading a book/listening to a song on the radio/watching a film at the cinema*?
 - a. How so?
10. Do you think using English in your free time like this has helped you improve your language skills?
 - a. Why/why not?
11. What do you think are the differences between learning in this way and learning in school?
 - a. How much of your English knowledge would you say comes from school, and how much from using the language in your free time? [estimate from 0–100%]
 - b. Why do you think that you learned more/less in your free time than in school?
12. Can you remember anything that you learnt using English outside of class that you don't think you could have learnt in class?
 - a. Can you give me some examples, like specific words or phrases; grammatical features; reading, writing, speaking, or listening abilities?
13. And my last question: If you could conduct a study on how young people your age use the internet to learn English, what would you want to know?

End of the Interview

Thank you for taking the time to participate in this discussion, your responses have been very helpful! That's all the questions I had for you. Is there anything else that you would like to add or something you would like to ask me? If you find that you have any questions for me later, please feel free to contact me by email at any time!

A.1.2 German

Begrüßung und allgemeine Informationen

Danke, dass ihr euch bereit erklärt habt, an meiner Studie teilzunehmen. Wie ihr schon von [eurem Lehrer/eurer Lehrerin] gehört habt, versuche ich mit meinem Projekt, mehr darüber herauszufinden, wie Schülerinnen und Schüler in eurem Alter Englisch im Alltag, außerhalb der Schule, verwenden. Die Ergebnisse meiner Studie werden hoffentlich dazu beitragen, zukünftigen Generationen von Schülerinnen und Schülern das Englisch Lernen weiter zu erleichtern.

Ich werde unser Gespräch mit dem Diktiergerät und meinem Telefon aufnehmen, damit ich nichts Wichtiges verpasse. Später werde ich alles transkribieren und die Aufnahme löschen, sodass danach alle Antworten vollkommen anonym sind. Alle Daten werden streng vertraulich behandelt und nicht an Dritte weitergegeben.

Ein paar Grundregeln:

1. Wenn ich euch irgendwelche Fragen stelle, die ihr nicht versteht oder von denen ihr möchtet, dass ich sie wiederhole, sagt mir bitte Bescheid.
2. Wenn es Fragen gibt, die ihr nicht beantworten möchtet, können wir sie einfach überspringen.
3. Es gibt für diese Fragen keine richtigen oder falschen Antworten – ich rechne damit, dass ihr alle verschiedene Erfahrungen und Meinungen habt. Bitte erzählt frei von euren Erfahrungen und Ansichten, auch wenn – bzw. gerade wenn – sie anders sind als die der Anderen.

Ihr solltet außerdem nicht das Gefühl haben, dass ihr immer mit mir sprechen müsst: ihr könnt einander antworten, zustimmen oder widersprechen. Ich bin nur hier, um euch Fragen zu stellen, zuzuhören und um darauf zu achten, dass jeder eine Chance hat, am Gespräch teilzunehmen.

Habt ihr noch Fragen an mich? Okay, dann lasst uns anfangen.

Fragen

1. Zuerst würde ich gerne ein wenig darüber hören, wie ihr in eurer Freizeit Englisch benutzt (also außerhalb der Schule oder eurer Hausaufgaben).
 - a. [Lest ihr zum Beispiel manchmal auf Englisch, hört Musik, schaut Filme, oder benutzt Englisch in sozialen Netzwerken wie Facebook und Twitter?]
 - b. Benutzt ihr zum Englisch lernen manchmal auch Lernprogramme im Internet, wie zum Beispiel Duolingo, Babbel oder Memrise, die euch nicht von euren Lehrern/Lehrerinnen vorgeschrieben wurden?
 - c. Will noch jemand etwas hinzufügen?

2. Was gefällt euch an diesen Aktivitäten?
 - a. [Es muss dabei nicht nur um die Sprache gehen. Zum Beispiel...]
3. Könnt ihr einschätzen, wie viel Zeit ihr pro Woche mit solchen Aktivitäten verbringt?
 - a. Ist das jede Woche mehr oder weniger gleich, oder variiert es?
 - b. Wenn ja, warum?
4. Von dem, was ihr bis jetzt gesagt habt, bekomme ich den Eindruck, dass ihr hauptsächlich auf Englisch *lest/schreibt*. Fallen euch Situationen ein, in denen ihr euch mit anderen auf Englisch austauscht (schriftlich oder gesprochen)?
 - a. [Schreibt ihr zum Beispiel Blogs, Facebook Status-Updates oder Tweets auf Englisch, oder schreibt ihr kurze Kommentare zu etwas, was andere gepostet haben (z.B. auf YouTube oder Facebook)?]
 - b. [Unterhaltet ihr euch mit anderen auf Englisch? Z.B. auf Facebook oder Skype?]
5. Warum macht ihr das — *Fernsehen/Lesen/im Internet surfen/etc.* — auf Englisch – statt auf Deutsch?
6. Ich würde gern von euch hören, wie gut ihr in diesen Situationen Englisch versteht. Das ist kein Test! Ich bin einfach daran interessiert zu hören, wie ihr eure eigenen Englischkenntnisse einschätzt. So kann ich mir ein besseres Bild davon machen, wie und warum ihr in eurer Freizeit Englisch benutzt.
 - a. Ist es euch wichtig, dass ihr alles (oder zumindest fast alles) versteht, wenn ihr auf Englisch *fernseht/lest/Spiele spielt/etc.*?
 - i. Warum/warum nicht?
 - b. Gibt es irgendwas, was ihr gerne auf Englisch machen würdet, was ihr aber nicht tut, weil ihr glaubt, dass ihr dazu nicht genug versteht?
7. Merkt ihr manchmal, dass euch ein einzelnes Wort oder eine grammatische Konstruktion bewusst auffällt, während ihr *Fernsehen schaut/Musik hört/lest/etc.*?
 - a. Warum/warum nicht?
 - b. Wie oft kommt das ungefähr vor?
8. Was macht ihr, wenn euch ein Wort oder eine grammatische Konstruktion ins Auge fällt?
 - a. Gibt es irgendetwas, was ihr tut, um das Wort oder das grammatische Merkmal zu lernen oder um es euch besser merken zu können?
9. Ist es für euch etwas Anderes, wenn ihr über euer Handy oder einen Computer im Internet surft, oder wenn ihr zum Beispiel ein *englisches Buch lest/ein englisches Lied im Radio hört/im Kino einen Film auf Englisch schaut*?
 - a. Wo liegen die Unterschiede?
10. Glaubt ihr, dass sich euer Englisch dadurch verbessert hat, dass ihr in es in eurer Freizeit so benutzt, wie wir hier besprochen haben?
 - a. Warum/warum nicht?

11. Was sind eurer Meinung nach die Unterschiede zwischen dieser Art zu lernen und dem Lernen in der Schule?
 - a. Wie viel eures Englischwissens, würdet ihr schätzen, kommt aus dem Schulunterricht und wie viel von Freizeitaktivitäten? [einschätzen von 0-100%]
 - b. Was sind Gründe dafür, dass ihr in der Schule/in der Freizeit mehr lernt?
12. Könnt ihr euch an etwas erinnern, das ihr in eurer Freizeit gelernt habt, was ihr nicht in der Schule gelernt hättet?
 - a. Nennt mir, wenn es geht, bitte einige Beispiele, z.B. bestimmte Wörter oder Redewendungen, Grammatik, Lesen, Hörverstehen, Schreibvermögen etc.
13. Und meine letzte Frage ist: Wenn ihr selber eine Studie darüber durchführen könntet, wie Jugendliche in eurem Alter im Internet Englisch lernen, was würdet ihr untersuchen wollen?

Gesprächsende

Danke, dass ihr euch die Zeit genommen habt, an meiner Studie teilzunehmen. Eure Antworten haben mir sehr weitergeholfen! Habt ihr noch Fragen an mich? Falls euch später noch Fragen einfallen, könnt ihr mir gerne eine E-Mail schicken!

A.2 Participant Questionnaire

The questionnaires were printed as 4-page, A5 brochures.

A.2.1 English



Thank you for taking part in my study! Please take some time to read these instructions carefully before you start filling in the questionnaire. Please answer the questions in this brochure as thoroughly as possible. Your answers will help me to better understand your contributions during our group discussion.

If something is unclear, please do not hesitate to ask. If you do not want to answer a certain question, you can skip it. However, the more questions you answer, the more useful your participation will be for our project. Feel free to write any comments in the margins of the questionnaire, if you think that there is any additional information that could be useful to us.

All of your data will be processed anonymously and confidentially.

Thank you again for your help! 😊

----- ✂ -----

Please write your name and your class below. I only need this information so that I can match your answers on this questionnaire with your contributions to our group discussion before. Don't worry, **your data will remain anonymous** — as soon as I have assigned you a participant ID, I will cut off this part of the questionnaire and throw it away.

Name:

Class:

1. **Gender:** ☐ male ☐ female ☐ other: _____

2. **How old are you?** _____

3. **What is your native language?** If you grew up bilingually, please list all languages that you learned from when you were young (under the age of 4).

a. ***If German is not your native language:*** How old were you when you started to learn German?

4. **How old were you when you started to learn English?**

If you are unsure, you can also write which year you were in when you first had English classes in school.

5. **Please use the following scale to rate your English proficiency in reading, writing, speaking, and listening.** Please tick only one of the circles in each row.

Reading	Not fluent, I only know a few simple sentences	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	Like a native speaker
Writing	Not fluent, I only know a few simple sentences	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	Like a native speaker
Speaking	Not fluent, I only know a few simple sentences	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	Like a native speaker
Listening	Not fluent, I only know a few simple sentences	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	Like a native speaker

6. Do you speak any other languages, besides German and English?

☐ No ☐ Yes

If you ticked "Yes", please write the languages in the following table and use the scale to rate your overall competency in each language. Please tick only one of the circles in each row.

Language:			
 _____	Not fluent, I only know a few simple sentences	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	Like a native speaker
_____	Not fluent, I only know a few simple sentences	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	Like a native speaker
_____	Not fluent, I only know a few simple sentences	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	Like a native speaker
_____	Not fluent, I only know a few simple sentences	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	Like a native speaker

7. Have you ever lived or been on holiday in an English-speaking country (longer than two weeks)? ☐ No ☐ Yes

If you ticked "Yes", please answer the following questions:

In which country?	When?	For how long?	Why? (e.g. holiday, language course, exchange, etc.)
 _____			

A.2.2 German



Danke, dass du an meiner Studie teilnimmst! Bitte nimm dir die Zeit, um diese Anleitung sorgfältig durchzulesen, bevor du beginnst, den Fragebogen auszufüllen.

Bitte beantworte die Fragen in diesem Heft so gründlich wie möglich. Sie helfen mir, deine Beiträge in unserem Gruppengespräch besser zu verstehen.

Wenn dir etwas unklar ist, wende dich bitte an mich. Wenn du eine Frage nicht beantworten möchtest, kannst du sie überspringen, aber je mehr Fragen du beantworten kannst, desto hilfreicher wird deine Teilnahme für unser Projekt sein. Du kannst gerne auch Anmerkungen auf den Fragebogen schreiben, wenn du meinst, dass zusätzliche Informationen für uns nützlich sein könnten.

Alle deine Daten werden anonym und vertraulich behandelt.

Noch einmal vielen Dank für deine Hilfe! 😊

----- ✂ -----

Bitte trage hier deinen Namen und deine Klasse ein. Diese Informationen brauche ich nur, um deine Antworten auf diesem Fragebogen und deine Beiträge in unserem Gespräch einander zuzuordnen. Keine Sorge, **deine Daten bleiben trotzdem anonym** – sobald ich dir eine Teilnehmernummer zugewiesen habe, schneide ich diesen Teil des Fragebogens ab und werfe ihn weg.

Name: _____

Klasse: _____

1. **Geschlecht:** ☐ männlich ☐ weiblich ☐ sonstige: _____

2. **Wie alt bist du?** _____

3. **Was ist deine Muttersprache?** Wenn du mehrsprachlich aufgewachsen bist, nenne bitte alle Sprachen, die du von klein auf (im Alter von unter 4 Jahren) gelernt hast.

a. **Falls Deutsch nicht deine Muttersprache ist: Wie alt warst du, als du angefangen hast, Deutsch zu lernen?**

4. **Wie alt warst du, als du angefangen hast, Englisch zu lernen?**

Wenn du dir nicht sicher bist, kannst du auch schreiben, in welcher Klasse du zuerst Englischunterricht hattest.

5. **Schätze mit Hilfe der Skala deine Englischkenntnisse beim Lesen, Schreiben, Sprechen, und Hörverstehen ein.** Kreuze jeweils nur einen Kreis an.

Lesen	Nicht fließend, ich kenne nur ein paar einfache Sätze	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	Wie ein Muttersprachler
Schreiben	Nicht fließend, ich kenne nur ein paar einfache Sätze	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	Wie ein Muttersprachler
Sprechen	Nicht fließend, ich kenne nur ein paar einfache Sätze	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	Wie ein Muttersprachler
Hörverstehen	Nicht fließend, ich kenne nur ein paar einfache Sätze	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	Wie ein Muttersprachler

6. Sprichst du außer Deutsch und Englisch noch andere Sprachen?

☐ Nein ☐ Ja

Wenn du "Ja" angekreuzt hast, schreibe die Sprachen in die untenstehende Tabelle und schätze anhand der Skala deine Gesamtkompetenz in jeder Sprache ein. Kreuze jeweils nur einen Kreis an.

Sprache:	Ich kenne nur ein paar einfache Sätze	O O O O O O O O	Ich beherrsche sie wie eine Muttersprache
 _____			

7. Hast du jemals in einem englischsprachigen Land gewohnt oder Urlaub gemacht (länger als 2 Wochen)? ☐ Nein ☐ Ja

Wenn du "Ja" angekreuzt hast, beantworte bitte die folgenden Fragen:

In welchem Land?	Wann?	Für wie lange?	Warum? (z.B. Urlaub, Sprachkurs, Schüleraustausch usw.)
 _____			

8. Hast du jemals in einem nicht englischsprachigen Land gewohnt oder Urlaub gemacht (länger als 2 Wochen)? ☐ Nein ☐ Ja

Wenn du "Ja" angekreuzt hast, beantworte bitte die folgenden Fragen:

In welchem Land?

Wann?

Für wie lange?

Warum?

(z.B. Urlaub, Sprachkurs,
Schüleraustausch usw.)



9. Wie oft du benutzt du die folgenden Geräte, um (außer für die Schule) auf das Internet zuzugreifen?

[illegible]

Appendix B. Methods Phase II and III

B.1 C-tests

Each version was prefaced with the same instructions in German, and an example text:

In den folgenden Texten fehlt bei einigen Wörtern ein Teil der Buchstaben. Bitte ergänze die Anfänge so, dass sinnvolle Wörter dabei herauskommen. Wichtig: Die Lücken sind immer gleich groß, egal wie viele Buchstaben fehlen. Abgekürzte Verbformen mit Apostroph (z.B. don't, isn't) zählen als ein Wort.

Hier ist ein Beispieltext:

I have a friend called Sam who hates December. Can y ou
guess w hy ? He's ta ll , quite la rge and h as a
b ig beard. H is boss alw ays asks h im to dr
ess up a s Santa Claus f or the chri stmas
party, a nd Sam does sn't like i t . He hates being
Santa Claus!

Versuche erst einmal den gesamten Text zu verstehen, bevor du etwas aufschreibst. Bitte schreibe deine Antworten auch dann in die Lücken, wenn du dir nicht ganz sicher bist, ob sie stimmen. Von dir wird kein perfektes Ergebnis erwartet — dieser Test ist selbst für Muttersprachler sehr schwierig!

The instructions translate to English as follows:

In the following texts, part of the letters is missing from some words. Please complete them to make meaningful words. Important: The gaps are all the same size, no matter how many letters are missing. Contracted verb forms with an apostrophe (e.g. don't, isn't) count as one word.

Here is an example text:

I have a friend called Sam who hates December. Can y ou
guess w hy ? He's ta ll , quite la rge and h as a
b ig beard. H is boss alw ays asks h im to dr
ess up a s Santa Claus f or the chri stmas
party, a nd Sam does sn't like i t . He hates being
Santa Claus!

Please try to understand the whole text before you write anything down. Please also fill in the gaps if you are not completely sure if your answer is correct. You are not expected to get a perfect score — this test can be very difficult even for native speakers!

B.1.1 Version 1

Includes the two texts that were excluded after piloting (Texts 3 and 4) and the advanced texts added after piloting (Texts 8 and 9). In the versions presented to the participants, texts were never split over two separate pages.

1. A Dog in Space

Laika was just a stray dog walking the streets of Moscow when she was chosen to become one of the most famous dogs in the world. In November 1957 she was sent into space and became the first living being to travel around the Earth. However, she did not come back to have her photograph taken as before the news. After just a few short hours in space she died aboard the spaceship Sputnik 2. It was her first and last flight but the start of many flights for humans.

2. Cheating in the Football World Cup

The World Cup is four weeks of footballers cheating from start to finish. How do they cheat? Well, there are a number of different ways. Firstly, they cheat 'dive'. This means falling over and pretending that they were pushed. Footballers also pretend to be kicked, for example. Then they roll over the ground and look like they are in real pain. Sometimes the referee gives a free-kick to the injured player's team. Then, after a few seconds the injured player gets up and is suddenly all right! The referee never says anything so the player gets away with it.

3. A Cab Ride with Consequences

A friend of a friend was in London for a conference. He took a taxi and there were no problems: the driver knew the way and was very cheerful. The trouble started when my friend realised that he had stupidly left his laptop on the back seat! In a panic, he went straightaway to the London and Finsbury office on the corner. He was over the moon to find the computer waiting for him, and then very surprised to learn that it was one of 2900 left in their taxis over the past six months!

4. A Young Chess Master

Truong Son learnt to play chess by watching his parents in their home. When he was first asked to play as well as they refused. But he insisted. So eventually they agreed and were astonished that he understood how to play. Why? Because little Truong Son was only three years old at that time! Chess is a difficult game to learn, with different pieces having different moves, but Truong Son knew them all. And, even more amazing, a month later he was already beating his parents at their favourite game!

5. Restaurant Reviews

The role of reviewers has been questioned in Australia recently, particularly the role of food critics, after a restaurant took a newspaper to court. They cla_____ that bec_____ the rev_____ in t_____ paper w_____ so b_____, people canc_____ reservations a_____ they lo_____ all o_____ their cust_____. The resta_____ had sp_____ a mil_____ dollars renov_____ the pl_____ before ope_____, but ju_____ a f_____ months af_____ the rev_____, it h_____ to cl_____ down. T_____ staff lo_____ their jobs, the owners lost their money and the restaurant lost its good reputation – all because of one review!

6. Birdsong in the City

Some birds are better at living in towns and adapting to a human world than others. For exa_____, in b_____ cities t_____ Blue Tit adj_____ its so_____ so th_____ it c_____ be he_____ more eas_____ over tra_____ noise. I_____ quiet str_____ the bi_____ have th_____ usual to_____, but i_____ busy, no_____ streets wi_____ a l_____ of tra_____ they si_____ higher s_____ that fem_____ can st_____ hear th_____. Adapting like this means they have a better chance of surviving. Unfortunately, other birds aren't so clever. The increasing noise and light from our cities means that they are unsuitable for many breeding birds.

7. The History of Tattoos

Although it's not certain exactly where tattooing started, there is lots of evidence that for millions of people throughout time it has been one of the most popular forms of permanent body art. It's be_____ suggested th_____ tattooing m_____ have sta_____ in Egypt bec_____ tattoos we_____ found o_____ Egyptian mum_____ from ab_____ 2000 BC. Ancient Maori war_____ also us_____ to pa_____ their fa_____ before a bat_____. This prov_____ camouflage a_____ protection aga_____ evil spi_____. They th_____ began t_____ make t_____ lines perm_____ rather th_____ reapplying th_____ each ti_____. In Borneo, men were also covered in images of plants and creatures. And in the South Seas Islands, particularly Samoa, tattooing continues to be an art form since it was first noted in the 19th century.

8. Tea and the British

There was a time in the 1960's and 1970's when some people believed that coffee would replace tea as Britain's favourite beverage but, in the end, that did not occur. To th_____ day, t_____ remains fir_____ anchored a_____ Britain's favo_____ drink, acco_____ for ov_____ 40% of a_____ drinks cons_____ in Britain wi_____ the exce_____ of wa_____. The ave_____ Briton ov_____ the a_____ of 10 dri_____ three a_____ a ha_____ cups o_____ tea p_____ day, o_____ 1,355

cu_____ per ye_____. This pu_____ Britain mi_____ ahead of any other country in the international league of tea-drinking nations!

9. America's Teenage Courts

The idea of "teen courts" has been around in the USA for many years: It was in the 1980s in Odessa that the Teen Court was first suggested. Realising th_____ many tee_____ offenders we_____ alienated b_____ a jus_____ system orga_____ and contr_____ by pe_____ of a diff_____ generation, the co_____ in Odessa dec_____ to l_____ offenders cho_____ to b_____ tried b_____ other teen_____. Teen cou_____ operate i_____ just t_____ same w_____ as a re_____ court, the ma_____ difference be_____ that t_____ only profes_____ is the judge and everyone else is a volunteer. While there is no possibility of an Odessa teen jury fining an offender or sending him or her to prison, there is a range of punishments available, including community service, driving classes, counselling and also jury service in the teen court.

B.1.2 Version 2

Includes the two texts that were excluded after piloting (Texts 3 and 4) and the advanced texts added after piloting (Texts 8 and 9). In the versions presented to the participants, texts were never split over two separate pages.

1. The Candle Ceremony

In Colombia, a very beautiful candle ceremony is a traditional part of every wedding. There a_____ three can_____, the br_____ lights t_____ candle o_____ her ri_____ and t_____ groom lig_____ the can_____ on h_____ left. Th_____, the n_____ husband a_____ wife li_____ the mid_____ candle. N_____ there a_____ three bur_____ candles. Th_____ the t_____ candles o_____ the le_____ and ri_____ are bl_____ out a_____ the middle candle is left to burn alone. This symbolises that the couple are now going to share every moment of their lives together.

2. The Most Loyal Dog in History

Hachiko was a large hunting dog and the faithful pet of a professor at Tokyo University. Every afte_____ after wo_____, the prof_____ came ba_____ to t_____ train sta_____ and Hachiko w_____ waiting pati_____ on t_____ platform f_____ him. Unfort_____ when Hachiko w_____ only 18 mon_____ old, t_____ professor di_____ before retu_____ home. O_____ course Hachiko d_____ not kn_____ this a_____ went t_____ the sta_____ to wa_____ for h_____ master w_____ did not come. Every day for the next ten years the dog continued to go to the station to meet the train. He died on the platform where he last saw the professor.

3. A Rude NYC Cabbie

Imagine how my sister felt when she went off for the first time to New York for a business trip. After a f_____ days s_____ was st_____ amazed a_____ how frie_____ everyone w_____ to h_____ until s_____ took a ta_____ back t_____ her ho_____ one ni_____. First s_____ gave ju_____ the na_____ of t_____ hotel, b_____ when t_____ driver as_____, she sa_____ the add_____ too. The dr_____ then wan_____ to kn_____ how t_____ get there but of course my sister didn't have a clue. Imagine how frightened she was when he started ranting and raving at her for not knowing the way! She said it was awful.

4. A Young Football Talent

Many little boys playing football in the street with friends dream of becoming another Pelé or Beckham. For Fredua Koranteng Adu, h_____ dream se_____ to b_____ coming tr_____. Already h_____ has foot_____ clubs li_____ Manchester United a_____ Chelsea sho_____ great inte_____. Why i_____ this s_____ remarkable? Bec_____ Freddy Adu i_____ only four_____ years o_____. He i_____ the teen_____ every ma_____ European te_____ wants t_____ sign. B_____ homework i_____ still impo_____ to Freddy! Wh_____ he was 11, Inter Milan wanted to take him to Italy, but his family wanted to keep him closer to continue his education.

5. Cat Survives Night Frozen to the Ground

Mrs Spears, an office cleaner from Chicago, arrived at work one cold, winter morning. The tempe_____ was ve_____ low: i_____ was begi_____ to sn_____. Hurrying t_____ go ins_____ and g_____ warm, s_____ was surp_____ to s_____ something ly_____ outside t_____ office do_____. She loo_____ closely a_____ was sho_____ to disc_____ it w_____ a c_____ frozen t_____ the gro_____. Mrs Spears real_____ it w_____ alive b_____ unable to move so she wrapped him up in her coat and brought him inside. Eventually the cat started to move and open his eyes. Now he lives in the office and has a new name: "Lucky"!

6. The History of the British Postal Service

In our world of email and text messaging it's almost impossible to imagine a time when you couldn't communicate to anyone in writing at all. Imagine n_____ knowing wh_____ was happ_____ to fri_____ and fam_____ who do_____ live wit_____ walking dist_____. Before 1635 th_____ was on_____ one per_____ who co_____ send a_____ receive let_____ in Bri_____ and th_____ was t_____ monarch - th_____ were on_____ carried t_____ and fr_____ the ro_____ court. I_____ was Ki_____ Charles the First w_____ allowed the use of the Royal Mail to members of the public and that's when the Post Office system was founded.

7. Harvesting Coffee Beans

Before becoming beans, the fruit of the coffee trees, known as a cherry, is harvested. The most common method is for the cherries to be hand-picked and dried in the sun on tables or on the ground. Once they are dried, all the outer layers are removed. Now the red cherries have become green beans. Now they are packed in 60kg bags and shipped abroad. Although the coffee is grown and harvested in the tropics, 70% of the total coffee production is imported by countries in the North who control the market.

8. Thanksgiving — A very American festival

The last weeks of the year are a festive time in most countries; but while Europeans just celebrate Christmas and the New Year, Americans begin their festive season about a month earlier. The holiday of Thanksgiving, celebrated on the fourth Thursday in November, is second in importance only to Christmas in the American calendar. Once a festival where whole communities celebrated together, Thanksgiving today is a day for families to celebrate. But apart from this, it has not changed greatly. Thanksgiving meal is still based, essentially, on native American foods that allowed early settlers to survive: turkey, corn, potatoes, squash and turkey— large birds that lived in the forests of North America.

9. Earthquakes in California: Waiting for "The Big One"

California is known all over the world for its size, its sun, and its surf, its glamour and its optimism — but it is known too as "Earthquake Country". It is a very vulnerable region where big devastating earthquakes have occurred in the past and could happen again. It's unlikely that a disaster on this scale or recent quakes in Italy or Japan could occur here. Nevertheless, Californians are worried about when the next quake will strike the state. Nobody knows for sure, but at all times California is on the alert. The earth is permanently monitored by universities and government research stations, where the instruments are constantly watched by highly-trained employees and volunteers from the California Office of Emergency Services.

B.2 Motivation and Attitudes Questionnaire

B.2.1 Overview of Motivation and Attitudes Constructs and Questionnaire Items

This table contains a complete list of the motivation and attitude constructs included in this study, their definitions, and the questionnaire items through which they were measured in the initial questionnaire (Phase II). Some of these items were excluded in the final questionnaire (Phase III, see below) or reassigned to a different construct as described in the main text (Chapter 7).

<i>Construct</i>	<i>Working definition</i>	<i>Questionnaire Items</i>
LANGUAGE LEARNING ORIENTATIONS		
<i>Intrinsic motivation</i>	Wanting to know English because one likes the language and finds the learning process enjoyable.	• I find English exciting. • I find learning English interesting. • I think learning English is really fun and I see it as a challenge. • I don't really like English. (R) • I feel frustrated whenever I try to speak and understand English. (R)
<i>Instrumental orientation: obligation</i>	Feeling externally pressure to learn English, e.g. by parents, through school.	• My parents want me to learn English. • I only learn English because I don't want to get bad marks. • I only want to learn enough English to pass my exams. • My friends don't really care about English. (R) • I would learn English even if others (e.g. my parents or friends) didn't think it was important. (R) • I would want to learn English, even if it wasn't a mandatory subject. (R)
<i>Instrumental orientation: added benefits</i>	Wanting to know English for the advantages it could offer in terms of future career, education, etc.	• English is important for my future career. • With good English skills, I will be able to advance more quickly in my future career. • With good English skills, I will be able to earn more money in future. • My English skills will help me get a better education (after school). • English won't play a big role in my further education (after school). (R) • Whether or not I will be successful in the future does not depend on my English skills. (R) • I don't think that English will be important in my future career. (R)

<i>Knowledge orientation</i>	Wanting to know English to help one to become better educated and pursue personal interest.	• My English skills help me to learn more things independently. • I want to know English to access information that is not available or difficult to access in German (e.g. news, wikis & other online texts). • It is not necessary to know English to stay informed of international news and events. • I want to know English so that I can acquire new ideas and broaden my horizon. • I want to know English because it helps me pursue my interests. • I don't need English to pursue my personal interests. (R)
<i>International orientation</i>	Wanting to know English to be able to interact with people from other cultural and language backgrounds.	• I want to know English because it helps you meet more people from other countries. • I want to know English to exchange ideas with people from other countries. • I want to know English so that I can learn more about the culture and life of people from other countries. • I don't think that knowing English helps you to better understand the culture and life of people from other countries. (R) • I don't think that English helps you meet more people from other countries. (R) • I don't think that knowing English makes you more "international". (R)
<i>Travel orientation</i>	Wanting to know English in order to travel or live abroad.	• I want to know English because I want to travel to other countries. • I want to know English because, in the future, I would like to live abroad for some time (e.g. working, studying). • I think it's important to know English when traveling to other countries. • I don't need to know English well because I don't like traveling to other countries. (R) • I don't need to know English well because I never plan on living abroad. (R)

L2 MOTIVATIONAL SELVES

<i>Ideal future L2 self</i>	The extent to which a learner imagines themselves as a successful user of their target language (English) in the future.	• I speak English fluently. • I can communicate effortlessly with others in English. • I can express my whole personality in English. • I can write effortlessly in English. • I can communicate successfully in English.
<i>Current L2 self</i>	The extent to which a learner currently perceives themselves to be a	• It's not a problem for me to use English. • I can read fluently in English. • I feel comfortable using English in different situations.

competent user of their
target language (English).

- Speaking English is part of who I am.
- I can express my thoughts and ideas without problems using English.

ATTITUDES TOWARDS LANGUAGE LEARNING

*Attitudes
towards
formal
language
learning*

Beliefs regarding the
relevance and effectiveness
of classes/formal learning
practices for language
learning.

- In general, I think that you effectively learn English in school.
- School classes are necessary to successfully learn English.
- Without school classes, I wouldn't know any English.
- I find the English that we learn in school more useful than what I encounter in my free time.
- When learning languages, it's important to know the grammar rules.
- When learning languages, it's important to explicitly practice vocabulary.
- What we do in English class doesn't really help you to learn the language. (R)
- You learn a lot *about* English in school, but not how to use the language. (R)
- 'School English' isn't useful to me in my daily life. (R)
- To me, English classes in school are a waste of time. (R)

*Attitudes
towards
informal
language
learning*

Beliefs regarding the
relevance and effectiveness
of informal language
learning practices.

- Most of my English knowledge comes from my free time.
- In general, I think that activities like watching TV, reading, listening to music, and playing video games can help you learn a language.
- I think that the internet can help you improve your English.
- The English I encounter in my free time is more useful to me than what we learn in school.
- In general, I prefer less formal methods for learning English.
- You only really learn a language through using it to communicate with others.
- The best way to learn a language is to use it as often as possible.
- I find English, the way I use it outside of school, *less important* than what we learn in school. (R)
- What I do in my free time does *not* really help me to learn English. (R)
- I don't like using English expressions that I learned in my free time, because you never know if they are linguistically correct. (R)

The questionnaire presented here is the final version, i.e. the one filled in by participants during Phase III. Previous versions can be supplied upon request. The questionnaires were printed as 12-page, A5 brochures.

B.2.2 English Questionnaire



Please take some time to read this information carefully **before** you start filling in the questionnaire.

If something is unclear, please do not hesitate to ask the project leader.

If you do not want to answer a certain question, you can skip it. However, the more questions you answer, the more useful your participation will be for this project.

Feel free to write any comments in the margins of the questionnaire, if you think that there is any additional information that could be useful to us.

You can still withdraw from participating in this project at any time.

All of your data will be treated confidentially and anonymously. They will not have any impact on your marks.

Thank you again for your help with this project!

Before you start on the next page,
please write down your participant ID here:

Section 1: Your language skills

1. What is your native language? ***If you grew up bilingually***, please list all languages that you learnt from before the age of 4.

1.a. ***If German is not your native language***: How old were you when you began learning German?

2. In which year at school did you first have English lessons?

3. Please use the scales below to estimate your proficiency in English. Please tick only ***one circle*** in each row.

Reading

I only know a few simple sentences

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

I know the language like a native speaker

Writing

I only know a few simple sentences

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

I know the language like a native speaker

Speaking

I only know a few simple sentences

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

I know the language like a native speaker

Listening

I only know a few simple sentences

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

I know the language like a native speaker

4. Please use the scales below to estimate how much each of the following factors contributed to your English knowledge. Please tick only ***one circle*** in each row.

English lessons and homework

Not at all

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

A lot

Independent learning (e.g. with apps, exercise books)

Not at all

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

A lot

Reading in English in my free time (e.g. books, magazines)

Not at all

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

A lot

Using English in daily life, outside of school lessons.

Not at all

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

A lot

Using English on holiday

Not at all ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ A lot

English-language films

Not at all ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ A lot

English-language serials and TV programmes

Not at all ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ A lot

English-language music

Not at all ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ A lot

English-language content on social media (e.g. Snapchat, YouTube)

Not at all ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ A lot

English-language computer games

Not at all ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ A lot

5. If you know any other languages besides German and English, please write them in the table below and use the scales to indicate your overall competence in each language.
Please tick only **one circle** in each row.



I only know a few simple sentences ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ I know the language like a native speaker

I only know a few simple sentences ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ I know the language like a native speaker

I only know a few simple sentences ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ I know the language like a native speaker

I only know a few simple sentences ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ I know the language like a native speaker

Section 2: Media

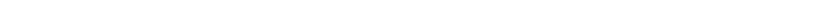
Here we want to find out what role different media play in your everyday life. Please use the scale below to indicate how often you use the following media:

Media	Rarely or never	A few times a month	A few times per week	Approximately once per day	Multiple times per day
Books					
German	☐	☐	☐	☐	☐
English	☐	☐	☐	☐	☐
Newspapers/magazines					
German	☐	☐	☐	☐	☐
English	☐	☐	☐	☐	☐
Music					
German	☐	☐	☐	☐	☐
English	☐	☐	☐	☐	☐
Films					
German	☐	☐	☐	☐	☐
English	☐	☐	☐	☐	☐
Serials/TV programs					
German	☐	☐	☐	☐	☐
English	☐	☐	☐	☐	☐
Social media (e.g. Snapchat, YouTube)					
German	☐	☐	☐	☐	☐
English	☐	☐	☐	☐	☐
Video games					
German	☐	☐	☐	☐	☐
English	☐	☐	☐	☐	☐
Other					
	☐	☐	☐	☐	☐
German	☐	☐	☐	☐	☐
English	☐	☐	☐	☐	☐

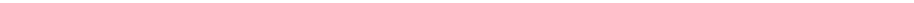
Section 3: Your attitudes about the English language

In this section, you will find a few statements about the English language. Some people agree with these statements, others think they are wrong. Please use the scales below to indicate how ***you personally*** feel about these statements. Please ***tick only one circle*** in each row.

I find English exciting.

Don't agree at all  Completely agree

I don't really like the English language.

Don't agree at all  Completely agree

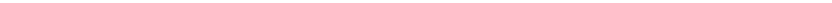
I want to know English because, in the future, I would like to live abroad for some time.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

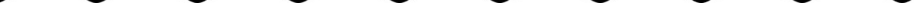
I don't need English to pursue my personal interests.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

I want to know English because it helps you meet more people from other countries.

Don't agree at all  Completely agree

I'm only learning English because I don't want to get bad marks.

Don't agree at all  Completely agree

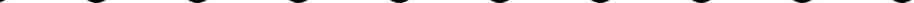
I don't think that English will be important in my future career.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

I don't need to know English well because I don't like traveling to other countries.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

My English skills will help me get a better education (after school).

Don't agree at all  Completely agree

I find learning English interesting.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

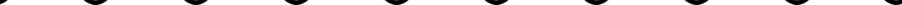
I would want to learn English, even if it wasn't a mandatory subject

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

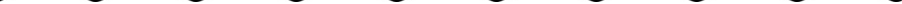
I want to know English so that I can learn more about the culture and life of people from other countries.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

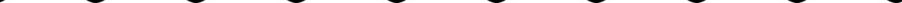
I want to know English because I want to travel to other countries.

Don't agree at all  Completely agree

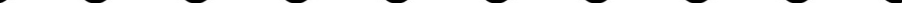
I think learning English is really fun and I see it as a challenge.

Don't agree at all  Completely agree

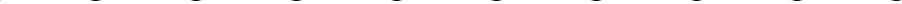
English won't play a big role in my further education (after school).

Don't agree at all  Completely agree

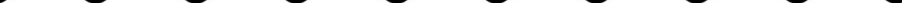
My English skills help me to learn more things independently.

Don't agree at all  Completely agree

I only want to learn enough English to pass my exams.

Don't agree at all  Completely agree

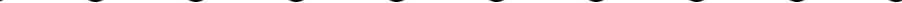
I don't really like English.

Don't agree at all  Completely agree

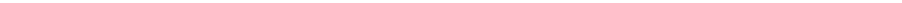
I want to know English because it helps me pursue my interests.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

I don't need to know English well because I never plan on living abroad.

Don't agree at all  Completely agree

I want to know English so that I can acquire new ideas and broaden my horizon.

Don't agree at all  Completely agree

Section 4: About you

1. Gender: ☐ male ☐ female ☐ other: _____

2. How old are you?

3. Do you have access to the internet outside of school? ☐ No ☐ Yes

If you ticked "yes", please indicate how often you use the following devices to access the internet in your free time:

Device	Rarely or never	A few times a month	A few times per week	Approximately once per day	Multiple times per day
Computer or laptop	☐	☐	☐	☐	☐
Tablet computer (e.g. iPad)	☐	☐	☐	☐	☐
Smartphone	☐	☐	☐	☐	☐
Other: 	☐	☐	☐	☐	☐
Other: _____	☐	☐	☐	☐	☐

4. Have you ever lived or been on holiday in an English-speaking country? ☐ No ☐ Yes

If you ticked "yes", please answer the following questions:

 In which country? For how long? Why? (e.g. holiday, language course, exchange, ...)

5. Have you ever lived or been on holiday in a **non-English-speaking** country (except Germany)?

☐ No ☐ Yes

If you ticked "yes", please answer the following questions:

In which country? **For how long?** **Why?** (e.g. holiday, language course, exchange, ...)



6. Here we want to find out about how you perceive yourself as an English-speaking person, *now and in the future*. Please indicate below each sentence how you personally feel about the statement. Please ***tick only one circle*** in each row. There are no right or wrong answers. It only matters that you tell us your personal opinion.

EXAMPLE: I speak English fluently.

I don't desire this at all	☐	☐	☐	☐	☐	☐	☐	☒	☐	☐	I desire this very much
I am still very far away from this	☐	☐	☒	☐	☐	☐	☐	☐	☐	☐	This is already completely true of me

I can communicate effortlessly with others in English.

I don't desire this at all	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	I desire this very much
I am still very far away from this	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	This is already completely true of me

I can express my whole personality in English.

I don't desire this at all	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	I desire this very much
I am still very far away from this	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	This is already completely true of me

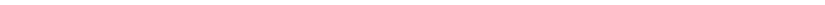
I can communicate successfully in English.

I don't desire this at all

I desire this very much

I am still very far away from this

This is already completely true of me

I am still very far away from this  This is already completely true of me

It's not a problem for me to use English.

I don't desire this at all ○○○○○○○○○○ I desire this very much

I am still very far away from this

This is already completely true of me

I feel comfortable using English in different situations.

I don't desire this at all ○○○○○○○○○○ I desire this very much

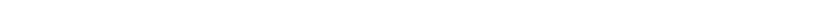
I am still very far away from this

This is already completely true of me

I can express my thoughts and ideas without problems using English.

I don't desire this at all ○○○○○○○○○○ I desire this very much

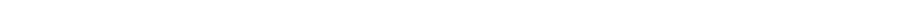
I am still very far away from this



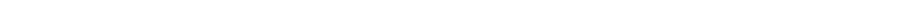
This is already completely true of me

7. Please indicate how much you agree with each of the three statements below. Please tick **only one circle** in each row.

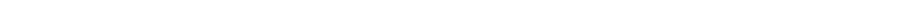
I think of myself as German.

Don't agree at all  Completely agree

I think of myself as European.

Don't agree at all  Completely agree

I think of myself as a global citizen.

Don't agree at all  Completely agree

Section 5: Your views on learning English

In this section, you will find a few statements about the English language. Some people agree with these statements, others think they are wrong. Please use the scales below to indicate how ***you personally*** feel about these statements. Please ***tick only one circle*** in each row.

I think that the internet can help you improve your English.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

What we do in English class doesn't really help you to learn the language.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

The English I encounter in my free time is more useful to me than what we learn in school.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

When learning languages, it's important to know the grammar rules.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

Most of my English knowledge comes from my free time.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

To me, English classes in school are a waste of time.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

What I do in my free time does not really help me to learn the language.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

When learning languages, it's important to explicitly practice vocabulary.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

I find English, the way I use it outside of school, less important than what we learn in school.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

In general, I think that you effectively learn English in school.

Don't agree at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Completely agree

Is there anything else that you would like to tell us? Then you can write it down here:

This is the end of the questionnaire. Please return the booklet to the researcher when you are done.

Thank you again for taking part in this study! 😊

B.2.3 German Questionnaire



Bitte lies dir diese Informationen sorgfältig durch, **bevor** Du anfängst, den Fragebogen auszufüllen.

Bitte beantworte die Fragen in diesem Heft so gründlich wie möglich. Wenn Dir etwas unklar ist, kannst Du dich an die Forscherin wenden.

Wenn Du eine Frage nicht beantworten möchtest, kannst Du sie überspringen. **Aber:** Je mehr Fragen Du beantwortest, desto hilfreicher wird Deine Teilnahme für dieses Projekt sein.

Du kannst gerne Anmerkungen auf den Fragebogen schreiben, wenn Du meinst, dass zusätzliche Informationen nützlich sein könnten.

Du kannst Deine Teilnahme an dieser Studie immer noch jederzeit abbrechen.

Alle Deine Daten werden anonym und vertraulich behandelt.

Vielen Dank für Deine Hilfe bei dem Projekt!

Bevor Du auf der nächsten Seite beginnst,
schreib bitte hier deine Teilnehmer-ID hin:

Abschnitt 1: Deine Sprachkenntnisse

1. Was ist Deine Muttersprache? **Wenn Du mehrsprachlich aufgewachsen bist**, nenne bitte alle Sprachen, die Du von klein auf (im Alter von unter 4 Jahren) gelernt hast.

1.a. **Wenn Deutsch nicht Deine Muttersprache ist:** Wie alt warst Du, als Du angefangen hast, Deutsch zu lernen?

2. In welcher Klasse hattest Du zuerst Englischunterricht?

3. Bitte schätze mit Hilfe der untenstehenden Skalen deine Englischkenntnisse ein. Kreuze jeweils nur **einen Kreis** an.

Lesen

Ich kenne
nur ein paar
einfache
Sätze

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Ich kenne die
Sprache wie ein
Muttersprachler

Schreiben

Ich kenne
nur ein paar
einfache
Sätze

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Ich kenne die
Sprache wie ein
Muttersprachler

Sprechen

Ich kenne
nur ein paar
einfache
Sätze

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Ich kenne die
Sprache wie ein
Muttersprachler

Hörverstehen

Ich kenne
nur ein paar
einfache
Sätze

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Ich kenne die
Sprache wie ein
Muttersprachler

4. Bitte schätze ab, wie sehr die folgenden Faktoren zu deinem heutigen Englischwissen beigetragen haben. Kreuze jeweils nur **einen Kreis** an.

Englischunterricht und Hausaufgaben

Überhaupt
nicht

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Sehr
stark

Selbstständiges Lernen (z.B. mit Apps, Übungsheften)

Überhaupt
nicht

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Sehr
stark

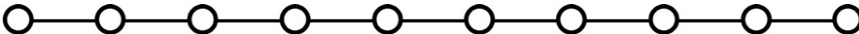
Englischsprachiges Lesen in der Freizeit (z.B. Bücher, Zeitschriften)

Überhaupt
nicht

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Sehr
stark

Englischer Sprachgebrauch im Alltag, außerhalb des Schulunterrichts

Überhaupt nicht  Sehr stark

Englischer Sprachgebrauch im Urlaub

Überhaupt nicht  Sehr stark

Englischsprachige Filme

Überhaupt nicht  Sehr stark

Englischsprachige Serien und Fernsehprogramme

Überhaupt nicht  Sehr stark

Englischsprachige Musik

Überhaupt nicht  Sehr stark

Englischsprachige Inhalte in sozialen Netzwerken (z.B. Snapchat, YouTube)

Überhaupt nicht  Sehr stark

Englischsprachige Computerspiele

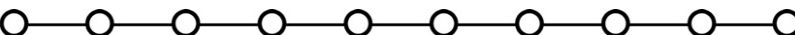
Überhaupt nicht  Sehr stark

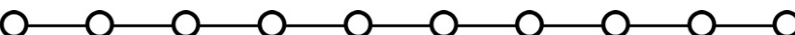
5. **Wenn Du außer Deutsch und Englisch noch andere Sprachen kennst**, schreibe die Sprachen in die untenstehende Tabelle und schätze anhand der Skala Deine **Gesamtkompetenz** in jeder Sprache ein. Kreuze jeweils **einen Kreis** an.



Ich kenne nur ein paar einfache Sätze  Ich kenne die Sprache wie ein Muttersprachler

Ich kenne nur ein paar einfache Sätze  Ich kenne die Sprache wie ein Muttersprachler

Ich kenne nur ein paar einfache Sätze  Ich kenne die Sprache wie ein Muttersprachler

Ich kenne nur ein paar einfache Sätze  Ich kenne die Sprache wie ein Muttersprachler

Abschnitt 2: Medien

Hier geht es darum, welche Rolle verschiedene Medien in deinem Alltag spielen.
Bitte gib an, wie oft Du auf die folgenden Medien zugreifst:

Medien	Selten oder nie	Ein paar Mal im Monat	Ein paar Mal in der Woche	Ungefähr einmal am Tag	Mehrmals täglich
Bücher					
deutschsprachig	☐	☐	☐	☐	☐
englischsprachig	☐	☐	☐	☐	☐
Zeitungen/Magazine					
deutschsprachig	☐	☐	☐	☐	☐
englischsprachig	☐	☐	☐	☐	☐
Musik					
deutschsprachig	☐	☐	☐	☐	☐
englischsprachig	☐	☐	☐	☐	☐
Filme					
deutschsprachig	☐	☐	☐	☐	☐
englischsprachig	☐	☐	☐	☐	☐
Serien/TV-Programme					
deutschsprachig	☐	☐	☐	☐	☐
englischsprachig	☐	☐	☐	☐	☐
Soziale Medien (z.B. Snapchat, YouTube)					
deutschsprachig	☐	☐	☐	☐	☐
englischsprachig	☐	☐	☐	☐	☐
Videospiele					
deutschsprachig	☐	☐	☐	☐	☐
englischsprachig	☐	☐	☐	☐	☐
Sonstige 	☐	☐	☐	☐	☐
deutschsprachig	☐	☐	☐	☐	☐
englischsprachig	☐	☐	☐	☐	☐

Abschnitt 3: Deine Einstellung zur englische Sprache

In diesem Abschnitt findest Du einige Aussagen über die englische Sprache. Manche Leute stimmen diesen Aussagen zu, andere halten sie für falsch. Bitte gib an, wie **Du persönlich** zu diesen Aussagen stehst. Kreuze jeweils **nur einen Kreis** an.

Ich finde Englisch spannend.

Stimme überhaupt nicht zu

Stimme voll und ganz zu

Ich mag die englische Sprache nicht wirklich.

Stimme überhaupt nicht zu

Stimme voll und ganz zu

Ich möchte Englisch können, weil ich einmal im Ausland leben möchte

Stimme überhaupt nicht zu

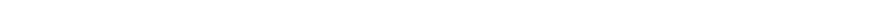
Stimme voll und ganz zu

Ich brauche Englisch nicht, um meinen persönlichen Interessen nachgehen zu können.

Stimme überhaupt nicht zu

Stimme voll und ganz zu

Ich möchte Englisch können, weil man damit mehr Menschen aus anderen Ländern treffen kann.

Stimme überhaupt nicht zu  Stimme voll und ganz zu

Ich lerne Englisch nur, weil ich keine schlechten Noten bekommen möchte.

Stimme überhaupt nicht zu

Stimme voll und ganz zu

Ich glaube nicht, dass Englisch für meinen zukünftigen Beruf wichtig sein wird.

Stimme überhaupt nicht zu

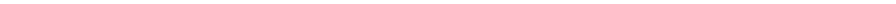
Stimme voll und ganz zu

Ich muss nicht gut Englisch können, denn ich reise nicht gerne in andere Länder.

Stimme überhaupt nicht zu

Stimme voll und ganz zu

Meine Englischkenntnisse werden mir zu einer guten weiteren Ausbildung verhelfen (nach der Schule).

Stimme überhaupt nicht zu  Stimme voll und ganz zu

Ich finde Englischlernen interessant.

Stimme überhaupt nicht zu

Stimme voll und ganz zu

Ich würde auch dann Englisch lernen wollen, wenn es kein Pflichtfach wäre.

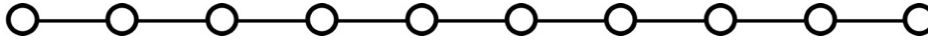
Stimme
überhaupt
nicht zu



Stimme
voll und
ganz zu

Ich möchte Englisch können, damit ich mehr über die Kultur und das Leben von Menschen aus anderen Ländern erfahre.

Stimme
überhaupt
nicht zu



Stimme
voll und
ganz zu

Ich möchte Englisch können, weil ich in andere Länder reisen will.

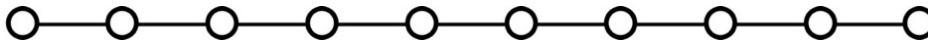
Stimme
überhaupt
nicht zu



Stimme
voll und
ganz zu

Ich habe echt Spaß am Englischlernen und sehe es als Herausforderung.

Stimme
überhaupt
nicht zu



Stimme
voll und
ganz zu

Englisch wird bei meiner weiteren Ausbildung (nach der Schule) keine große Rolle spielen.

Stimme
überhaupt
nicht zu



Stimme
voll und
ganz zu

Meine Englischkenntnisse helfen mir, selbstständig mehr zu lernen.

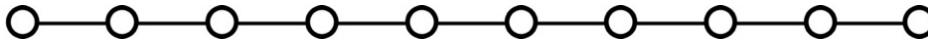
Stimme
überhaupt
nicht zu



Stimme
voll und
ganz zu

Ich möchte nur so viel Englisch lernen, wie ich brauche, um die Klassenarbeiten zu bestehen.

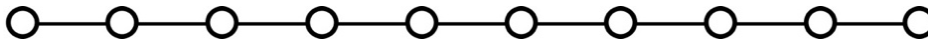
Stimme
überhaupt
nicht zu



Stimme
voll und
ganz zu

Ich mag Englisch nicht wirklich.

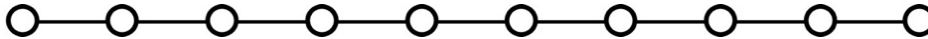
Stimme
überhaupt
nicht zu



Stimme
voll und
ganz zu

Ich möchte Englisch können, weil es mir dabei hilft, meinen Interessen nachzugehen.

Stimme
überhaupt
nicht zu



Stimme
voll und
ganz zu

Ich muss nicht gut Englisch können, weil ich nicht vorhabe, jemals im Ausland zu leben.

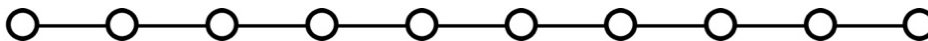
Stimme
überhaupt
nicht zu



Stimme
voll und
ganz zu

Ich möchte Englisch können, um neue Ideen zu bekommen und meinen Horizont zu erweitern.

Stimme
überhaupt
nicht zu



Stimme
voll und
ganz zu

5. Hast Du jemals in einem **nicht englischsprachigen** Land gewohnt ☐ Nein ☐ Ja
oder Urlaub gemacht?

Wenn Du „Ja“ angekreuzt hast, beantworte bitte die folgenden Fragen:

In welchem Land? **Für wie lange?** **Warum?** (z.B. Urlaub, Sprachkurs, Austausch, ...)



6. Hier geht es darum, wie Du Dich *jetzt und in der Zukunft* als „englischsprachige Person“ wahrnimmst. Gib bitte unter jedem Satz an, wie Du persönlich zu den einzelnen Aussagen stehst. Bitte **kreuze auf jeder Skala genau einen Kreis an**. Es gibt keine richtigen oder falschen Antworten. Wichtig ist nur, dass Du Deine persönliche Meinung angibst.

BEISPIEL: Ich spreche fließend Englisch.

Das wünsche ich mir gar nicht	○	○	○	○	○	○	○	✗	○	○	Das wünsche ich mir sehr
Davon bin ich noch sehr weit entfernt	○	○	✗	○	○	○	○	○	○	○	Trifft jetzt schon voll auf mich zu

Ich kann mich mühelos mit anderen auf Englisch verständigen.

Das wünsche ich mir gar nicht	○	○	○	○	○	○	○	○	○	○	Das wünsche ich mir sehr
Davon bin ich noch sehr weit entfernt	○	○	○	○	○	○	○	○	○	○	Trifft jetzt schon voll auf mich zu

Ich kann auf Englisch meine ganze Persönlichkeit zum Ausdruck bringen.

Das wünsche ich mir gar nicht	○	○	○	○	○	○	○	○	○	○	Das wünsche ich mir sehr
Davon bin ich noch sehr weit entfernt	○	○	○	○	○	○	○	○	○	○	Trifft jetzt schon voll auf mich zu

Ich kann erfolgreich auf Englisch kommunizieren.Das wünsche
ich mir gar
nicht
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Das
wünsche ich
mir sehrDavon bin ich
noch sehr
weit entfernt
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Trifft jetzt
schon voll
auf mich zu**Es ist für mich kein Problem, Englisch zu verwenden.**Das wünsche
ich mir gar
nicht
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Das
wünsche ich
mir sehrDavon bin ich
noch sehr
weit entfernt
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Trifft jetzt
schon voll
auf mich zu**Ich fühle mich wohl dabei, in verschiedenen Situationen Englisch zu benutzen.**Das wünsche
ich mir gar
nicht
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Das
wünsche ich
mir sehrDavon bin ich
noch sehr
weit entfernt
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Trifft jetzt
schon voll
auf mich zu**Ich kann meine Gedanken und Ideen mühelos auf Englisch rüberbringen.**Das wünsche
ich mir gar
nicht
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Das
wünsche ich
mir sehrDavon bin ich
noch sehr
weit entfernt
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Trifft jetzt
schon voll
auf mich zu

7. Bitte gib hier an, wie sehr Du den folgenden drei Aussagen zustimmst.

Kreuze jeweils **nur einen Kreis** an.**Ich betrachte mich selbst als deutsch.**Stimme
überhaupt
nicht zu
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Stimme
sehr
stark zu**Ich betrachte mich selbst als europäisch.**Stimme
überhaupt
nicht zu
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Stimme
sehr
stark zu**Ich betrachte mich selbst als Weltbürger/in.**Stimme
überhaupt
nicht zu
☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
Stimme
sehr
stark zu

Abschnitt 5: Deine Ansichten zum Englischlernen

Hier findest Du einige Aussagen übers Englischlernen. Manche Leute stimmen diesen Aussagen zu, andere halten sie für falsch. Bitte benutze die untenstehende Skala, um anzugeben, wie **Du persönlich** zu diesen Aussagen stehst. Bitte kreuze jeweils **nur einen Kreis** an.

Ich glaube, dass das Internet einem helfen kann, sein Englisch zu verbessern.

Stimme überhaupt nicht zu

Stimme sehr stark zu

Was wir im Englischunterricht machen hilft einem wirklich, die Sprache zu lernen.

Stimme überhaupt nicht zu

Stimme sehr stark zu

Das Englisch, das mir in meiner Freizeit begegnet, finde ich nützlicher als das, was wir in der Schule lernen.

Stimme überhaupt nicht zu

Stimme sehr stark zu

Beim Sprachen lernen ist es wichtig, die grammatischen Regeln zu kennen.

Stimme überhaupt nicht zu

Stimme sehr stark zu

Der Großteil meiner Englischkenntnisse kommt aus meiner Freizeit.

Stimme überhaupt nicht zu

Stimme sehr stark zu

Englischunterricht in der Schule ist für mich Zeitverschwendung.

Stimme überhaupt nicht zu

Stimme sehr stark zu

Was ich in meiner Freizeit mache, hilft mir nicht wirklich dabei, Englisch zu lernen.

Stimme überhaupt nicht zu

Stimme sehr stark zu

Beim Sprachenlernen ist es wichtig, Vokabeln ausdrücklich zu üben.

Stimme überhaupt nicht zu

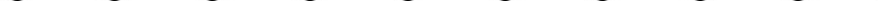
Stimme sehr stark zu

Englisch so wie ich es außerhalb der Schule benutze, finde ich weniger wichtig als das, was wir in der Schule lernen.

Stimme überhaupt nicht zu

Stimme sehr stark zu

Ganz allgemein denke ich, dass man im Schulunterricht effektiv Englisch lernt.

Stimme überhaupt nicht zu  Stimme sehr stark zu

**Ist Dir beim Ausfüllen des Fragebogens noch etwas
eingefallen, was Du uns gerne mitteilen möchtest?
Dann kannst Du es hier aufschreiben.**

Dies ist das Ende des Fragebogens. Gib das Heft bitte an die
Forscherin zurück, wenn Du fertig bist.

Noch einmal vielen Dank für Deine Teilnahme! 😊

B.3 Informal Second Language Learning Diary

B.3.1 Overview of Engagement Constructs and Survey Items

This table contains a list of the engagement constructs in this study, their definitions, and the items through which they were measured in the diary survey (Phase II). Some of these items were excluded in the final survey (Phase III).

<i>Construct</i>	<i>Working definition</i>	<i>Survey Items</i>
<i>Behavioural engagement</i>	Observable qualities of engagement with ISLPs, including frequency, duration, diversity of practices.	<i>Type of activity & language skill used:</i> • ‘What did you do?’ • Various follow-up questions per activity <i>Duration:</i> • ‘How much time did you approximately spend on this activity?’ <i>Frequency:</i> • Number/distribution of diary entries.
<i>Cognitive engagement</i>	The learner’s cognitive state while engaging in ISLPs, including the extent to which they are focusing on the activity itself vs. extraneous stimuli (focus on task at hand) and on understanding the meaning of the text they are processing (focus on meaning).	• ‘I felt [not at all – completely] absorbed in this activity.’ • ‘During this activity, I [did not concentrate at all – concentrated hard] on it.’ • ‘It was [very easy – very difficult] to keep my mind on the activity.’ • ‘All in all, I was [not at all – completely] focused on the content.’
<i>Affective engagement</i>	The learner’s emotional states while engaging in ISLPs, including enjoyment, interest, and challenge-skill balance.	• ‘All in all, this activity was [boring – entertaining]’ • ‘I [really liked – really did not like] this activity.’ • ‘The content was [not at all interesting – very interesting] to me.’ • ‘During this activity, I felt [confident – overwhelmed].’
<i>Linguistic engagement</i>	The extent to which the learner is focused on processing linguistic features and improving their language skills.	• ‘I was [not at all – completely] focused on trying to understand every single word.’ • ‘All in all, I was [not at all – completely] focused on the language.’ • ‘All in all, I thought [very little – a lot] about the language during this activity.’ • ‘This activity was [very – not at all] beneficial to my English learning.’

B.3.2 English Instructions

Please read these instructions carefully!

In this study, we want to find out how pupils in Germany come into contact with English outside of school. For this reason, we ask you to keep a record **every day for seven days**, starting today, of how you use English in your free time. You can do this by filling in a short online survey **every time** you do one of the following:

- Look at English content on social media (e.g. Instagram, Snapchat, Facebook, etc.)
- Watch a film, a series, a TV programme, a YouTube video, etc. in English
- Play a video game in English
- Listen to music in English
- Read a text in English (e.g. a book, a text on the internet)
- Write something in English yourself
- Hear English spoken (e.g. a conversation or a radio programme)
- Speak English yourself

Please fill in the survey **from the start every time** you have done one of these things in English. It will take only 1 or 2 minutes and you can do it on your smartphone or a computer.

If you didn't do any of these things on a given day, please fill in the survey anyway and choose the option "I didn't use English outside of school today". At the end of the week, you should have submitted *at the very least* 7 entries.

To keep your data anonymous, we will assign you a participant ID that you have to provide every time you fill in the survey. Your participant ID is:

A01

You can find the survey at: [Hyperlink](#)

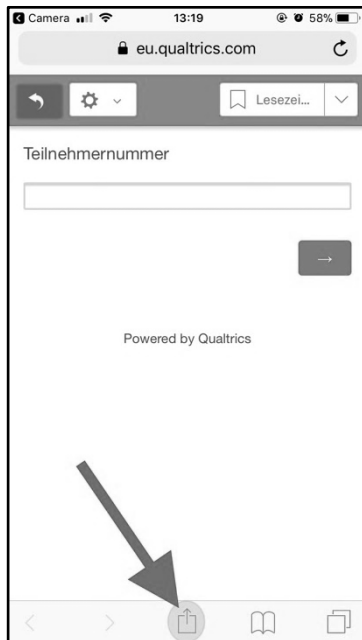
If you have a smartphone, the easiest way for you to fill in the survey will be to save it as an app on your home screen, so that you have easy access to it. The next pages contain instructions for how you can do this (on an iPhone or Android phone).

If you have any further questions about the survey, you can ask your teacher. You can also send us an email. You can reach the research you have already met under: henriette.arndt@education.ox.ac.uk

Thank you again for your help! 😊

Instructions for saving the survey as an app (iPhone)

1. Pull up the survey ([Hyperlink](#)) in Safari and tap the arrow symbol at the bottom of the screen.



2. In the menu, tap the button “Add to Home Screen”. If you do not see the button right away, you may have to swipe left in the lower part of the menu until it appears.



3. Now you will see a text box, where you can give the app a name (e.g. “Digital English”). When you are done, tap “Add”.



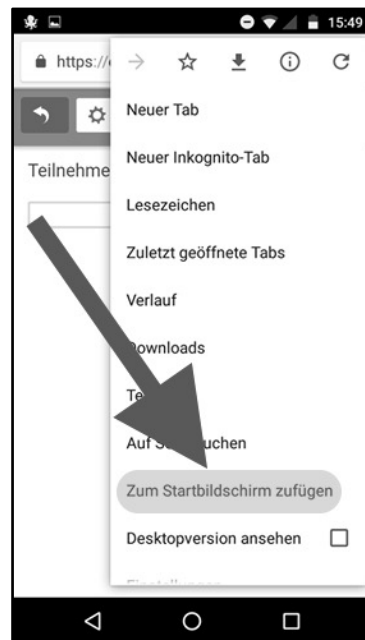
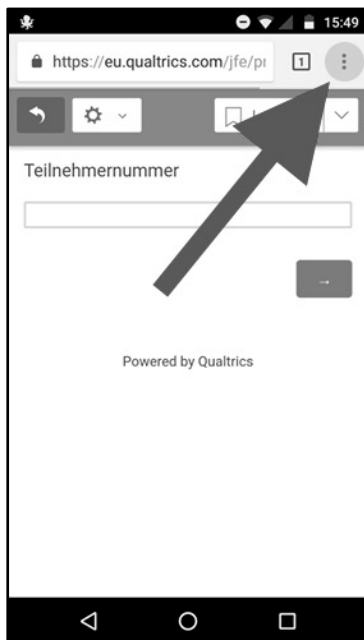
4. Now the app will appear on your Home Screen. When you tap it, the survey will open in your browser.



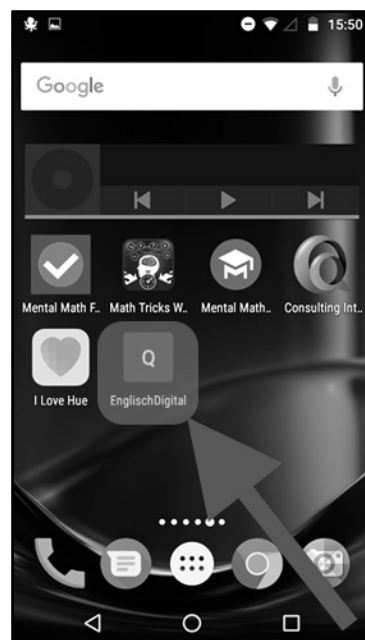
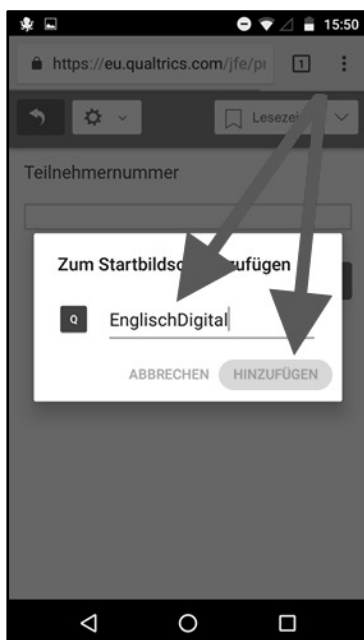
Please don't forget to delete the app at the end of the study!

Instructions for saving the survey as an app (Android)

1. Pull up the survey ([Hyperlink](#)) in Chrome and tap the three-dot symbol at the top of the screen.
2. In the menu, tap “Add to Home Screen”.



3. Now you will see a text box, where you can give the app a name (e.g. “Digital English”). When you are done, tap “Add”.



Please don't forget to delete the app at the end of the study!

B.3.3 German Instructions

Bitte lies Dir diese Anleitung sorgfältig durch!

In dieser Studie möchten wir herausfinden, wie Schülerinnen und Schüler in Deutschland außerhalb der Schule mit Englisch in Kontakt kommen. Deswegen bitten wir Dich, ab heute **sieben Tage lang jeden Tag** festzuhalten, wie Du in Deiner Freizeit Englisch benutzt. Das funktioniert, indem Du **jedes Mal** eine kurze Online-Umfrage ausfüllst, wenn Du eines der folgenden Dinge gemacht hast:

- Englischsprachige Inhalte auf sozialen Netzwerken angesehen (z.B. Instagram, Snapchat, Facebook usw.)
- Einen Film, eine Serie, ein TV-Programm, ein YouTube-Video usw. auf Englisch geguckt
- Ein Videospiel auf Englisch gespielt
- Musik auf Englisch gehört
- Einen Text auf Englisch gelesen (z.B. ein Buch, einen Text im Internet)
- Selber etwas auf Englisch geschrieben
- Englisch gesprochen gehört (z.B. ein Gespräch oder ein Radioprogramm)
- Selbst Englisch gesprochen

Bitte fülle die Umfrage **jedes Mal neu** aus, wenn Du eine der oben genannten Sachen auf Englisch gemacht hast. Das dauert nur 1-2 Minuten und geht ganz einfach auf Deinem Smartphone oder am Computer.

Wenn du an einem Tag keine dieser Dinge gemacht hast, fülle die Umfrage trotzdem aus und wähle die Option „Ich habe Englisch heute außerhalb der Schule nicht benutzt“. Am Ende der Woche solltest du also *mindestens* 7 Einträge übermittelt haben.

Damit Deine Daten anonym bleiben, bekommst Du eine Teilnehmer-ID, die Du jedes Mal bei der Umfrage mit angeben musst. Deine Teilnehmer-ID ist:

A01

Die Umfrage findest Du unter [Hyperlink](#)

Wenn Du ein Smartphone hast, kannst Du die Umfrage am besten als App auf Deinem Startbildschirm speichern, sodass Du sie schnell und einfach wiederfinden kannst. Anleitungen dazu (für iPhone und Android) findest Du auf den nächsten Seiten.

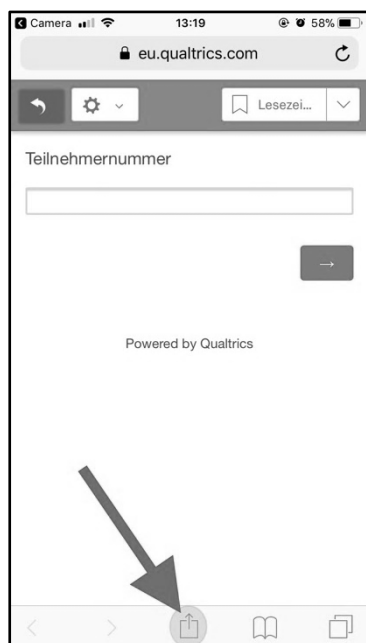
Wenn Du noch Fragen zu der Umfrage hast, wende Dich bitte an Deine/n Englischlehrer/in.

Du kannst uns auch eine E-Mail schreiben. Die Forscherin, die Du schon kennen gelernt hast, erreichst du unter: henriette.arndt@education.ox.ac.uk

Noch einmal vielen Dank für Deine Mithilfe! 😊

Anleitung zum Speichern der Umfrage als App (iPhone)

1. Rufe die Umfrage ([Hyperlink](#)) über Safari auf und tippe auf das Pfeilsymbol in der Mitte.
2. Tippe in dem Menü auf den Button „Zum Home-Bildschirm“. Eventuell siehst du diesen Button nicht sofort, sondern musst im unteren Menüabschnitt nach links wischen bis er erscheint.



3. Jetzt erscheint ein Textfeld, wo Du der App einen Namen geben kannst (z.B. „EnglischDigital“). Tippe danach auf „Hinzufügen“.

4. Jetzt erscheint die App auf Deinem Startbildschirm. Wenn Du sie antippst, öffnet sich die Umfrage in Deinem Browser.

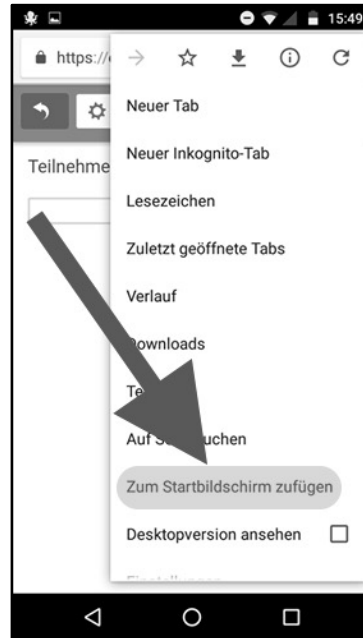
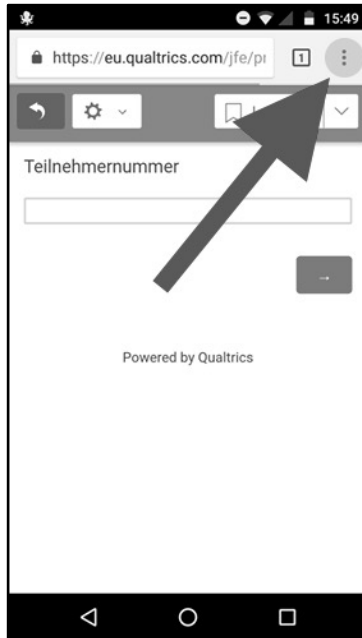


Bitte vergiss nicht, diese App am Ende der Studie wieder zu löschen!

Anleitung zum Speichern der Umfrage als App (Android)

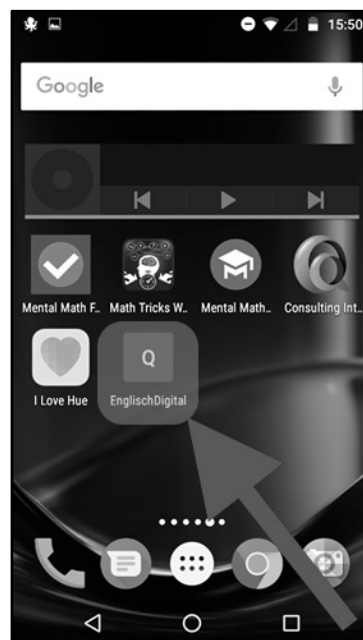
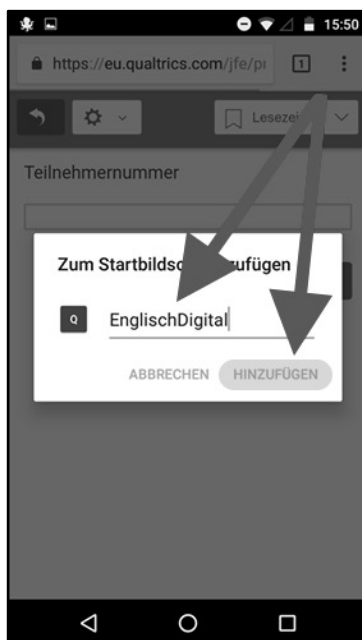
1. Rufe die Umfrage ([Hyperlink](#)) über Chrome auf und tippe auf das Drei-Punkte Symbol oben rechts.

2. Tippe in dem Menü auf „Zum Startbildschirm zufügen“.



3. Jetzt erscheint ein Textfeld, wo Du der App einen Namen geben kannst (z.B. „EnglischDigital“). Tippe danach auf „Hinzufügen“.

4. Jetzt erscheint die App auf Deinem Startbildschirm. Wenn Du sie antippst, öffnet sich die Umfrage in Deinem Browser.



Bitte vergiss nicht, diese App am Ende der Studie wieder zu löschen!

B.3.4 English List of Questions

Below overviews, in English and in German, of the questions contained in the online survey that constitutes the engagement diary. It should be noted that the participants were not presented with all this information at any one time. Rather, page skip logic within the survey was employed, so that the participants only saw the questions relevant to the activity they were recording (page 2). An example of a path that a participant might take through the survey is presented in the screenshots after the German instructions. These screenshots also show what the survey looked like on the participants' phones.

Page 1: Identification

Participant ID: _____

Page 2: Activity

What did you do?

- ☐ Look at English contents on Social Networking Sites
- ☐ Listen to music in English
- ☐ Play a video game in English (e.g. on the computer or a games console)
- ☐ Watch a film, TV programme, etc. in English
- ☐ Read a text in English
- ☐ Write a text in English
- ☐ Listen to spoken English (e.g. a radio programme or a conversation other people had)
- ☐ Have a conversation in English yourself
- ☐ Other – Please explain what you did: _____

Page 3: Follow-up questions, depending on the chosen activity

3.1. Social Networks

Which Social Networking Site did you visit?

- ☐ Instagram
- ☐ YouTube
- ☐ Snapchat
- ☐ Facebook
- ☐ Twitter
- ☐ Other – Please explain which website/app you used: _____

What did you do on Instagram? (You can select more than one answer)

- ☐ Read photo captions in English
- ☐ Watch videos in English
- ☐ Read English-language comments by others
- ☐ Write your own comment or text in English
- ☐ Other – Please explain what you did _____

What did you do on YouTube? (You can select more than one answer)

- ☐ Watch videos in English
- ☐ Read English subtitles
- ☐ Read English-language comments by others
- ☐ Write your own comment or text in English
- ☐ Other – Please explain what you did _____

What did you do on Snapchat? (You can select more than one answer)

- ☐ Read photo captions in English
- ☐ Watch videos in English
- ☐ Read articles in English in the Discovery section
- ☐ Read English-language comments or messages by others
- ☐ Write your own comment or message in English
- ☐ Other – Please explain what you did _____

What did you do on Facebook? (You can select more than one answer)

- ☐ Watch videos in English
- ☐ Read English-language photo captions, comments or status updates
- ☐ Read linked texts in English (e.g. newspaper articles, blog posts)
- ☐ Write your own English photo captions, comments or status updates
- ☐ Used messenger to **write** with someone in English
- ☐ Used messenger to **speak** with someone in English
- ☐ Other: Please explain what you did _____

What did you do on Twitter? (You can select more than one answer)

- ☐ Read tweets in English
- ☐ Write your own tweets in English
- ☐ Read linked texts in English (e.g. newspaper articles, blog posts)
- ☐ Watch linked videos in English
- ☐ Other: Please explain what you did _____

Other: What exactly did you do? (You can select more than one answer)

- ☐ Watch videos in English
- ☐ Read texts or comments in English
- ☐ Write texts or comments in English
- ☐ Have a written conversation with someone in English
- ☐ Speak with someone in English
- ☐ Other: Please explain what you did _____

3.2. Music

Did you...

- ☐ ... have the music running in the background
- ☐ ... listen attentively to the music

Did you... (You can select multiple answers or leave this question blank)

- ☐ ... read the lyrics
- ☐ ... sing along

3.3. Films, TV programmes, etc.

What did you watch?

- ☐ An English-language film
- ☐ An English-language TV programme
- ☐ An online video (e.g. on YouTube)
 - *Links to questions about YouTube under 3.2.*
- ☐ Other: Please explain what you watched _____

How did you watch the film/TV programme?

- ☐ On the internet (e.g. on Netflix, Amazon Prime)
- ☐ On TV
- ☐ On DVD
- ☐ I downloaded the film/programme

Did you use subtitles?

- ☐ Yes, English subtitles
- ☐ Yes, German subtitles
- ☐ No subtitles

3.4. Video games

What kind of game did you play?

- ☐ Single-player game
- ☐ Multi-player game

→ Single-player game: What did you do while playing?

(You can select more than one answer)

- ☐ Listen to spoken English
- ☐ Read English
- ☐ Write English (e.g. text commands)

→ Multi-player game: What did you do while playing?

(You can select more than one answer)

- ☐ Listen to spoken English
- ☐ Read English
- ☐ Write English (e.g. text commands)
- ☐ Have a **written** conversation with others
- ☐ Have a **spoken** conversation with others

3.5. Reading

What did you read?

- ☐ An English-language book
- ☐ An article in a book or magazine (online or on paper)
- ☐ Other texts on the Internet (e.g. blog posts, Wikipedia, fan fiction)
- ☐ Other: Please explain what you read _____

3.6. Writing

What did you write?

- ☐ A short text or message (e.g. Whatsapp or Facebook Messenger)
- ☐ A longer text (e.g. an E-Mail, a blog post)
- ☐ Other: Please explain what you wrote _____

3.7. Spoken English

In what context did you hear English spoken?

- ☐ An English-language audiobook
- ☐ An English-language radio programme
- ☐ An English-language podcast
- ☐ A conversation that others were having in English
- ☐ Other: Please explain what you listened to _____

3.8. Speak English yourself

How did the conversation take place?

- ☐ In person
- ☐ On the phone, Skype, etc.

→ In person: Who were you speaking to in English?

- ☐ Someone I know personally (e.g. relatives, friends)
- ☐ A stranger (e.g. a tourist in your city)

→ Phone, Skype: Who were you speaking to in English?

- ☐ Someone I know personally (e.g. relatives, friends)
- ☐ Someone I have never met in person (e.g. an online friend)

Page 4: Duration

How much time did you approximately spend on this activity?

_____ hours _____ minutes

Page 5: Engagement

All in all, this activity was ...

entertaining  boring

I felt ... absorbed in this activity

completely  not at all

The contents was ... to me.

very interesting  not at all interesting

I was ... focused on trying to understand every single word.

not at all  completely

During this activity, I ...
 concentrated hard  did not concentrate at
 all

All in all I was ... focused on the content.
 not at all  completely

All in all I was ... focused on the language.
 not at all  completely

All in all, I thought ... about the language during this activity.
 very little  a lot

During this activity, I felt...
 confident  overwhelmed

It took ... to keep my mind on what I was doing.
 no effort at all  a lot of effort

This activity was ... beneficial to my English learning.
 very  not at all

All in all, I ... this activity.
 really liked  really did not like

B.3.5 German List of Questions

1. Seite: Identifikation

Teilnehmernummer: _____

2. Seite: Aktivität

Was hast du gemacht?

- ☐ Englische Inhalte auf sozialen Netzwerken angesehen
- ☐ Musik auf Englisch gehört
- ☐ Ein Videospiel auf Englisch gespielt (z.B. am Computer oder an einer Konsole)
- ☐ Einen Film, eine Fernsehserie o.ä. auf Englisch angesehen
- ☐ Einen Text auf Englisch gelesen
- ☐ Selbst Englisch geschrieben
- ☐ Gesprochenes Englisch gehört (z.B. ein Radioprogramm oder ein Gespräch überhört)
- ☐ Selbst ein Gespräch auf Englisch geführt
- ☐ Sonstiges – Bitte erkläre, was du gemacht hast: _____

3. Seite: Weiterführende Fragen abhängig von der Aktivität

3.1. Soziale Netzwerke

Welches soziale Netzwerk hast du besucht?

- ☐ Instagram
- ☐ YouTube
- ☐ Snapchat
- ☐ Facebook
- ☐ Twitter
- ☐ Sonstiges – Bitte erkläre, welche Webseite/App du besucht hast: _____

Was hast du auf Instagram gemacht? (Du kannst mehrere Antworten ankreuzen)

- ☐ Bildunterschriften auf Englisch gelesen
- ☐ Englischsprachige Videos angeguckt
- ☐ Englischsprachige Kommentare von anderen gelesen
- ☐ Selber einen Kommentar oder Text auf Englisch geschrieben
- ☐ Sonstiges – Bitte erkläre, was du gemacht hast: _____

Was hast du auf YouTube gemacht? (Du kannst mehrere Antworten ankreuzen)

- ☐ Englischsprachige Videos angeguckt
- ☐ Englischsprachige Untertitel gelesen
- ☐ Englischsprachige Kommentare von anderen gelesen
- ☐ Selber einen Kommentar auf Englisch geschrieben
- ☐ Sonstiges – Bitte erkläre, was du gemacht hast: _____

Was hast du auf Snapchat gemacht? (Du kannst mehrere Antworten ankreuzen)

- ☐ Bildunterschriften auf Englisch gelesen
- ☐ Englischsprachige Videos angeguckt
- ☐ Über die Discover-Funktion englischsprachige Artikel gelesen
- ☐ Englischsprachige Kommentare oder Nachrichten von anderen gelesen
- ☐ Selber eine Nachricht oder einen Kommentar auf Englisch geschrieben
- ☐ Sonstiges – Bitte erkläre, was du gemacht hast: _____

Was hast du auf Facebook gemacht? (Du kannst mehrere Antworten ankreuzen)

- ☐ Englischsprachige Videos angeguckt
- ☐ Englischsprachige Bildunterschriften, Kommentare oder Status-Updates anderer gelesen
- ☐ Verlinkte Texte (z.B. Zeitungsartikel oder Blog-Posts) auf Englisch gelesen
- ☐ Selber Bildunterschriften, Kommentare oder Status-Updates auf Englisch geschrieben
- ☐ Messenger genutzt, um mit jemandem auf Englisch zu **schreiben**
- ☐ Messenger genutzt, um mit jemandem auf Englisch zu **sprechen**
- ☐ Sonstiges – Bitte erkläre, was du gemacht hast: _____

Was hast du auf Twitter gemacht? (Du kannst mehrere Antworten ankreuzen)

- ☐ Englischsprachige Tweets gelesen
- ☐ Eigene Tweets auf Englisch geschrieben
- ☐ Verlinkte Texte (z.B. Zeitungsartikel oder Blog-Posts) auf Englisch gelesen
- ☐ Verlinkte Videos auf Englisch angesehen
- ☐ Sonstiges – Bitte erkläre, was du gemacht hast: _____

Sonstige: Was genau hast du gemacht? (Du kannst mehrere Antworten ankreuzen)

- ☐ Englischsprachige Videos angesehen
- ☐ Englischsprachige Texte oder Kommentare gelesen
- ☐ Selber Texte oder Kommentare auf Englisch geschrieben
- ☐ Dir auf Englisch mit jemandem geschrieben
- ☐ Dich auf Englisch mit jemandem unterhalten
- ☐ Sonstiges: Bitte erkläre, was du gemacht hast _____

3.2. Musik

Hast du die Musik...

- ☐ ... im Hintergrund laufen lassen
- ☐ ... aufmerksam angehört

Hast du dabei... (Du kannst mehrere oder gar keine Antworten ankreuzen)

- ☐ ... die Texte mitgelesen
- ☐ ... mitgesungen

3.3. Film, Fernsehserie, o.ä.

Was hast du geschaut?

- ☐ Einen englischsprachigen Film
- ☐ Ein englischsprachiges Fernsehprogramm
- ☐ Ein Onlinevideo (z.B. auf YouTube)
 - Verlinkt zu Fragen über YouTube unter 3.2.
- ☐ Sonstiges – Bitte erkläre, was du geschaut hast: _____

Wie hast du den Film/das Fernsehprogramm geschaut?

- ☐ Im Internet (z.B. auf Netflix, Amazon Prime)
- ☐ Im Fernsehen
- ☐ Auf DVD
- ☐ Ich habe den Film/das Programm heruntergeladen

Hast du dabei Untertitel angehabt?

- ☐ Ja, englische Untertitel
- ☐ Ja, deutsche Untertitel
- ☐ Nein, keine Untertitel

3.4. Videospiele

Welche Art von Spiel hast du gespielt?

- ☐ Singleplayer-Spiel
- ☐ Multiplayer-Spiel

→ Singleplayer-Spiel: Was hast du beim Spielen gemacht?

(Du kannst mehrere Antworten ankreuzen)

- ☐ Englisch gehört
- ☐ Englisch gelesen
- ☐ Englisch geschrieben (z.B. Textbefehle)

→ Multiplayer-Spiel: Was hast du beim Spielen gemacht?

(Du kannst mehrere Antworten ankreuzen)

- ☐ Englisch gehört
- ☐ Englisch gelesen
- ☐ Englisch geschrieben (z.B. Textbefehle)
- ☐ Dir mit anderen auf Englisch **geschrieben**
- ☐ Englisch mit anderen **gesprochen**

3.5. Lesen

Was hast du gelesen?

- ☐ Ein englischsprachiges Buch
 - ☐ Einen Artikel in einer Zeitung oder einem Magazin (im Internet oder auf Papier)
 - ☐ Andere Texte im Internet (z.B. Blog-Posts, Wikipedia, Fan Fiction)
 - ☐ Sonstiges – Bitte erkläre, was du gelesen hast:
-

3.6. Schreiben

Was hast du geschrieben?

- ☐ Einen kurzen Text oder eine Nachricht (z.B. Whatsapp oder Facebook Messenger)
 - ☐ Einen längeren Text (z.B. eine E-Mail, einen Blog-Post)
 - ☐ Sonstiges – Bitte erkläre, was du geschrieben hast:
-

3.7. Gesprochenes Englisch

In welchem Zusammenhang hast du Englisch gesprochen gehört?

- ☐ Ein englisches Hörbuch
- ☐ Ein englischsprachiges Radioprogramm
- ☐ Einen englischen Podcast
- ☐ Ein Gespräch, was andere auf Englisch geführt haben
- ☐ Sonstiges – Bitte erkläre, wo du Englisch gesprochen gehört hast:

3.8. Selbst Englisch sprechen

Wie hat die Unterhaltung stattgefunden?

- ☐ Persönlich
- ☐ Übers Telefon, Skype, o.ä.

→ Persönlich: Mit wem hast du dich auf Englisch unterhalten?

- ☐ Jemand, den ich persönlich kenne (z.B. Verwandte, Freunde)
- ☐ Ein/e Fremde/r (z.B. Touristen in deiner Stadt)

→ Telefon, Skype: Mit wem hast du dich auf Englisch unterhalten?

- ☐ Jemand, den ich persönlich kenne (z.B. Verwandte, Freunde)
- ☐ Jemand, den ich noch nie persönlich getroffen habe (z.B. Online-Freunde)

4. Seite: Dauer

Wie lange hast du ungefähr mit dieser Aktivität verbracht?

_____ Stunden _____ Minuten

5. Seite: Engagement

Insgesamt fand ich die Aktivität ...

unterhaltsam  langweilig

Ich habe mich ... in die Aktivität vertieft gefühlt.

überhaupt nicht  vollkommen

Der Inhalt war für mich ...

sehr interessant  überhaupt nicht
interessant

Ich war ... darauf bedacht, jedes einzelne Wort zu verstehen.

überhaupt nicht  sehr stark

Während der Aktivität habe ich mich ...

voll darauf konzentriert  gar nicht darauf
konzentriert

Insgesamt war ich ... auf den **Inhalt** fokussiert.

überhaupt nicht  vollkommen

Insgesamt war ich ... auf die **Sprache** fokussiert.

überhaupt nicht  vollkommen

Insgesamt habe ich bei der Aktivität ... über die Sprache nachgedacht.

sehr wenig  sehr viel

Während dieser Aktivität habe ich mich ... gefühlt.

selbstsicher  überwältigt

Mit meinen Gedanken bei der Aktivität zu bleiben war für mich ...

überhaupt nicht anstrengend  sehr anstrengend

Diese Aktivität hat ... dazu beigetragen, meine Englischkenntnisse zu verbessern.

überhaupt nicht  sehr stark

Alles in allem hat mir die Aktivität...

sehr gefallen  überhaupt nicht gefallen

B.3.6 Screenshots of the Online Survey

Teilnehmer-ID (Zahl + Buchstabe, z.B. A01)

→

Powered by Qualtrics

Was hast du gemacht?

Englische Inhalte auf sozialen Netzwerken angesehen

Musik auf Englisch gehört

Einen Film, eine Serie, ein TV-Programm o.ä. auf Englisch geschaut

Ein Videospiel auf Englisch gespielt (z.B. am Computer oder einer Spielekonsole)

Einen Text auf Englisch gelesen

Selbst Englisch geschrieben

Gesprochenes Englisch gehört (z.B. ein Radioprogramm, oder ein Gespräch überhört)

Selbst ein Gespräch auf Englisch

Was hast du dir angesehen?

Einen englischsprachigen Film

Eine englischsprachige Serie oder ein TV-Programm

Ein Onlinevideo (z.B. auf YouTube)

Sonstiges: Bitte erkläre, was du geschaut hast

← →

Powered by Qualtrics

Wie hast du dir den Film/die Serie/das Programm angesehen?

Im Internet (z.B. auf Netflix, Amazon Prime)

Im Fernsehen

Auf DVD

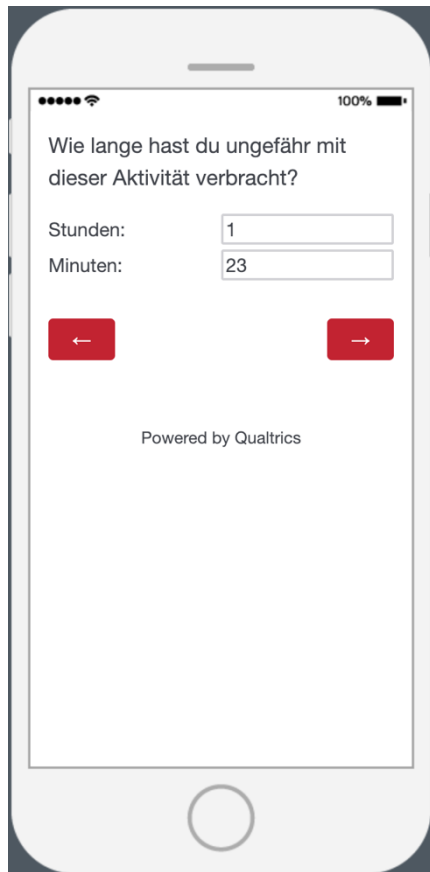
Ich habe den Film/die Serie/das Programm runtergeladen

Hast du dabei Untertitel angehab?

Ja, englische Untertitel

Ja, deutsche Untertitel

Nein, keine Untertitel



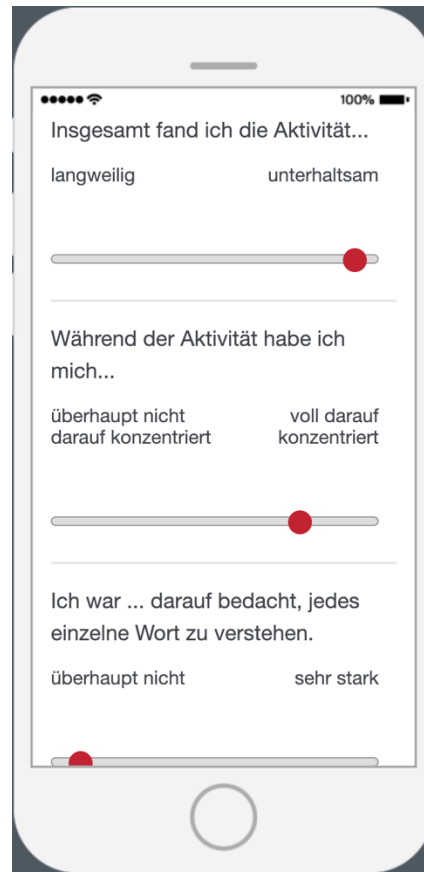
Wie lange hast du ungefähr mit dieser Aktivität verbracht?

Stunden:

Minuten:

← →

Powered by Qualtrics



Insgesamt fand ich die Aktivität...

langweilig unterhaltsam

Während der Aktivität habe ich mich...

überhaupt nicht darauf konzentriert voll darauf konzentriert

Ich war ... darauf bedacht, jedes einzelne Wort zu verstehen.

überhaupt nicht sehr stark

Appendix C. Findings Phase I

C.1 Qualitative Coding Framework

ENGAGEMENT	
Affective engagement	The participants' emotional states while engaging in ISLPs, including their levels of interest and enjoyment and how much they liked the activities.
Enjoyment	Having fun with an informal, deriving pleasure from it.
Interest	Wanting to find out more about or engage with something, being enthusiastic about the subject of an activity.
Topics	Subjects that in which the participants were interested (e.g. sports, music, fashion, beauty,).
Negative	Negative emotions associated with certain ISLPs, e.g. embarrassment, frustration.
Behavioural engagement	Observable qualities of engagement in ISLPs, including frequency, duration, variety of practices, and language skills used.
Frequency	How often participants engaged in particular ISLPs.
Events	Activities in which the participants engaged very occasionally (signalled, for example, by the use of phrases such as 'once' and 'right now').
Practices	Activities in which the participants engage regularly or often (e.g. once per month, multiple times per week).
Time spent	The time that participants estimated they spent on ISLPs per day or per week (individual practices or total).
Check-in	Activities in which participants engaged often, but briefly (mostly social media or short written communication).
Intense engagement	Activities in which participants engaged very frequently, for long amounts of time (e.g. playing video games) or a background activity (e.g. playing music while also doing something else).
Fluctuations	Reasons for why participants spent more/less time on ISLPs during some periods of time.
Types of activities	
Books	
Films	
Learning Apps	
Magazines	
Music	
Lyrics	
Singing along	
Serials	
Social interactions	
Holidays	
Online only	Interactions with people that the participants had met online/didn't know personally.
Online: personal acquaintances	Interactions with people with whom the participants had personal connections (e.g. friends, extended family).
Speaking	
Writing	
Social media	
Facebook	
Instagram	
Twitter	
YouTube	

Video games	
Multi-player games	
Single-player games	
Other activities	
Cognitive engagement	The learner's cognitive state while engaging in ISLPs, including the extent to which they were focused on the activity itself vs. extraneous stimuli (focus on task at hand) and on understanding the meaning of the text they were processing (focus on meaning).
Comprehension	The extent to which the learners were able to understand the general meaning of the task they engaged with and how easy/difficult it was for them to understand.
Focus	The extent to which participants paid attention only to the informal activity itself vs. extraneous stimuli.
Linguistic engagement	The extent to which participants focused on processing linguistic features and improving their language skills.
Strategies for comprehension	Conscious cognitive or behavioural strategies that the participants employed purely to increase their understanding of the overall meaning.
Strategies for learning	Conscious cognitive or behavioural strategies that the participants employed with the intention of increasing the linguistic gains from engaging in ISLPs.
Motivation	
Language learning motivation	The reasons for wanting to know/learn English.
International orientation	Wanting to know English to be able to interact with people from other cultural and language backgrounds.
Instrumental benefit	Wanting to know English for the advantages it could offer in terms of future career, education, etc.
Instrumental obligation	Feeling externally pressured to learn English, e.g. by parents, through school.
Intrinsic motivation	Wanting to know English because of an inherent liking of the language and finding the process of learning it enjoyable.
Knowledge orientation	Wanting to know English to pursue personal interests and learn new things.
Travel orientation	Wanting to know English in order to travel or live abroad.
Situated motivation	The reasons for (not) engaging in ISLPs (in-the-moment, context-dependent).
Bad quality translation	Dissatisfaction with the quality of the German translation (content that was originally in English, e.g. books, serials).
Difficulty with comprehension	Problems with not being able to comprehend the language.
Incidental exposure	Informal engagement due to presence of the language in the environment (e.g. someone else playing an English movie).
Intrinsic language preference	Having an inherent preference for the English language.
Lack of access	Not having easy access to English-language content (particularly on-demand video streaming, e.g. Netflix).
Popular culture	Engaging with popular culture that happens to be in English, rather than intentionally choosing English-language culture.
ATTITUDES	
Affective states	Subjective experience of informal vs. formal learning.
Informal learning	Beliefs regarding the authenticity and effectiveness of informal language learning practices.
Incidental learning	Informal learning as a by-product of informal practices.
Intentionality	Informal learning as an intentional act.
Formal learning	Beliefs regarding the authenticity and effectiveness of formal language learning practices.
Learning gains	Language learning outcomes of informal/formal learning as perceived by the students.

Confidence
Everyday vs. 'school' English
Colloquial language as it is spoken by natives/proficiency language users vs. more formal varieties taught in school
Fluency/automatization
Grammar and syntax
Listening comprehension
Pronunciation
Reading comprehension
Speaking skills
Vocabulary

Appendix D. Findings Phase II

D.1 C-tests Preliminary Analyses: Normality and Outliers

Table D.1: C-tests descriptive statistics. Percentile scores by text and version

Item	n	min	max	M	SD	Z _{skewness}	Z _{kurtosis}
VERSION 1							
Text 1	77	22.00	100.00	67.82	18.35	- 1.58	- .79
Text 2	77	28.00	100.00	66.68	18.98	- .52	- 2.26
Text 3	77	24.00	90.00	63.48	16.33	- 3.33	1.06
Text 4	77	20.00	100.00	70.23	14.99	- 1.07	.44
Text 5	77	8.00	100.00	59.44	22.70	- .48	- .96
Text 6	77	4.00	96.00	42.78	24.28	2.04	- 1.06
Text 7	77	16.00	100.00	50.75	23.12	- 1.45	-1.11
Total	77	30.00	96.00	60.17	16.36	- 1.30	- .46
VERSION 2							
Text 1	79	4.00	100.00	62.66	22.25	- 1.85	- .29
Text 2	79	28.00	100.00	79.16	16.38	- 3.85	1.75
Text 3	79	26.00	100.00	81.46	15.19	- 5.71	5.98
Text 4	79	4.00	100.00	70.10	19.54	- 3.93	2.17
Text 5	79	22.00	100.00	68.89	24.02	- 1.48	- 2.25
Text 6	79	12.00	100.00	57.72	25.65	.21	- 2.06
Text 7	79	0.00	100.00	44.0	27.54	.62	- 1.84
Total	79	28.00	96.29	66.29	17.24	- 1.40	- 1.23

The numbers in bold mark where the skewness and kurtosis exceeded the range for a 'normal' distribution.

D.2 C-tests: Lexical Diversity and Readability Statistics

Table D.2: C-tests readability statistics by version

<i>Text</i>	<i>Number of words in the text</i>	<i>Type-Token Ratio – TTR</i> (high score → high complexity)	<i>Flesch Reading Ease Score</i> (high score → low complexity)	<i>Gunning Fog Index</i> (high score → high complexity)	<i>Dale-Chall Score</i> (high score → high complexity)
VERSION 1					
<i>Text 1</i>	95	0.67	75.0	8.8	5.9
<i>Text 2</i>	119	0.69	84.5	7.4	5.0
<i>Text 3</i>	95	0.64	76.2	10.1	6.1
<i>Text 4</i>	115	0.67	75.6	9.6	7.0
<i>Text 5</i>	110	0.62	56.3	14.7	6.8
<i>Text 6</i>	128	0.75	54.6	13.9	7.8
<i>Text 7</i>	105	0.70	68.4	10.5	7.7
<i>Text 8</i>	99	0.67	63.4	13.2	7.6
<i>Text 9</i>	136	0.62	32.5	22.2	7.2
VERSION 2					
<i>Text 1</i>	91	0.56	67.8	8.0	5.6
<i>Text 2</i>	109	0.61	62.7	7.9	6.0
<i>Text 3</i>	110	0.63	85.6	7.4	5.8
<i>Text 4</i>	111	0.75	60.1	8.3	7.2
<i>Text 5</i>	109	0.65	54.6	8.8	6.6
<i>Text 6</i>	105	0.66	56.6	14.2	7.0
<i>Text 7</i>	95	0.66	56.0	12.9	7.2
<i>Text 8</i>	118	0.69	38.1	17.9	7.3
<i>Text 9</i>	128	0.68	47.5	15.7	8.6

These statistics were calculated using www.textinspector.com (TTR, Flesch Reading Ease Score, Gunning Fog Index) and www.datayze.com/readability-analyzer (Dale-Chall Score).

Brief explanations of the measurements:

- **Type-Token Ratio** (Templin, 1957): This measure of lexical variation initially comes from research about child language development. It is calculated by dividing the total number of *different* words ('types') in a text by the sum of all running words ('tokens'). Higher values indicate more lexical variation, and thus more complexity.
- **Flesch Reading Ease Score** (Flesch, 1948): This readability score is calculated based on the number or total words, total sentences, and total syllables in a text. Higher scores indicate a *less* complex text: Scores can range from 1 to 100 and a conversion table is available to interpret these numbers. For example, a score of 50-60 is thought to correspond to a 10-12th grade native-speaker reading level; a score below 30 indicates college-level complexity.

- **Gunning Fog Index** (Gunning, 1952): This is a readability index similar to the Flesch Reading Ease Score in that it is based on sentence length and syllable counts, namely the percentage of 'foggy' words (those that contain three or more syllables). A score of 5 indicates that a document is 'readable', 10 is 'hard', and 20 is 'very difficult'.
- **Dale-Chall Score** (Dale & Chall, 1948): This is one of the most accurate readability metrics because instead of relying on syllable counts, this score draws on a list of the most common 3,000 words in English to identify words that are easier to understand. Scores range from 0 to 10. Values below 5 indicate a primary-school reading level, while scores above 9 correspond to college-level complexity.

Figure D.1: Lexical diversity & readability scores for C-test Version 1

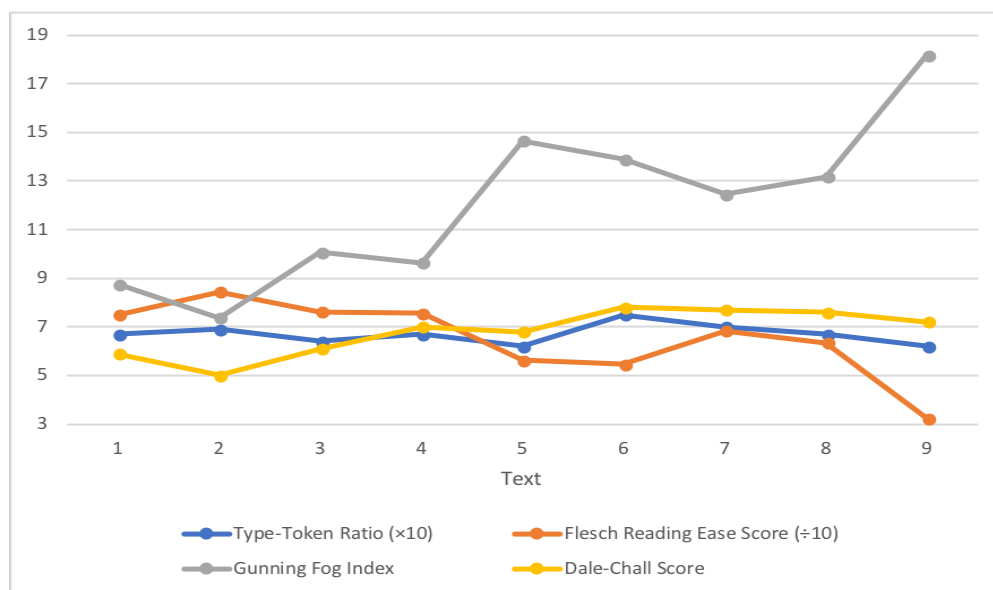
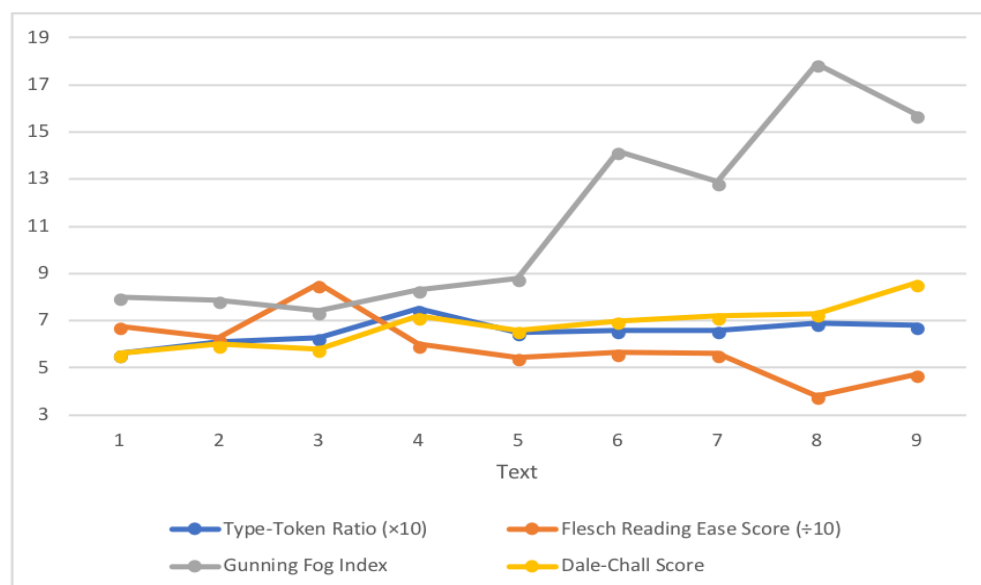


Figure D.2: Lexical diversity & readability scores for C-test Version 2



D.3 Exploratory Factor Analyses

D.3.1 C-Tests

Table D.3: C-tests Exploratory Factor Analysis. Summary of results and descriptive statistics after excluding Texts 1–4 in various combinations

<i>Excluded</i>	<i>Version 1</i>	<i>Version 2</i>
<i>Texts 1 & 2</i>	• Good: All communalities > .60 • Total variance explained: 71.3% • 80% non-redundant residuals > .05 • Total scores: $M = 57.4, SD = 17.9$	• Acceptable: Only one communality < .60 • Total variance explained: 67.5% • 70% non-redundant residuals > .05 • Total scores: $M = 64.7, SD = 18.7$
<i>Texts 1 & 3</i>	• Acceptable: Only one communality < .60 • Total variance explained: 71.9% • 50% non-redundant residuals > .05 • Total scores: $M = 58.0, SD = 17.8$	• Good: All communalities > .60 • Total variance explained: 67.2% • 70% non-redundant residuals > .05 • Total scores: $M = 64.0, SD = 18.6$
<i>Texts 1 & 4</i>	• Good: All communalities > .60 • Total variance explained: 73.9% • 70% non-redundant residuals > .05 • Total scores: $M = 56.6, SD = 18.1$	• Good: All communalities > .60 • Total variance explained: 69.4% • 60% non-redundant residuals > .05 • Total scores: $M = 66.2, SD = 18.2$
<i>Texts 2 & 3</i>	• Good: All communalities > .60 • Total variance explained: 71.9% • 70% non-redundant residuals > .05 • Total scores: $M = 57.4, SD = 17.9$	• Poor: Two communalities < .60 • Total variance explained: 66.6% • 70% non-redundant residuals > .05 • Total scores: $M = 64.8, SD = 18.7$
<i>Texts 2 & 4</i>	• Acceptable: Only one communality < .60 • Total variance explained: 71.1% • 70% non-redundant residuals > .05 • Total scores: $M = 57.0, SD = 18.3$	• Acceptable: Only one communality < .60 • Total variance explained: 68.3% • 70% non-redundant residuals > .05 • Total scores: $M = 63.3, SD = 19.1$
<i>Texts 3 & 4</i>	• Good: All communalities > .60 • Total variance explained: 71.1% • 70% non-redundant residuals > .05 • Total scores: $M = 57.5, SD = 18.1$	• Good: All communalities > .60 • Total variance explained: 68.4% • 60% non-redundant residuals > .05 • Total scores: $M = 62.5, SD = 19.2$

Based purely on the results from these Exploratory Factor Analyses and the reliability analyses, excluding Texts 1 and 4 seemed to be the best option for both versions. However, as one of the goals of these exclusions was to balance the difficulty of the two tests, the researcher also inspected the effect that the exclusions would have on the total mean scores on each version. In the end, it was decided to exclude Text 3 and 4 in each version, as this option seemed to provide the best compromise between optimising the results of the EFA and reliability tests and minimising the difference between the two version means.

D.3.2 Language Learning Orientations

Table D.4: Language learning orientations Exploratory Factor Analysis. Communalities in the four- and five-factor model

Item	Four-factor model	Five-factor model
<i>I find English exciting.</i>	.65	.81
<i>I find learning English interesting.</i>	.83	.85
<i>I think learning English is really fun and I see it as a challenge.</i>	.65	.72
<i>I don't really like English. (R)</i>	.82	.82
<i>I think that knowing English will help me advance in my future career.</i>	.48	.64
<i>I don't think that English will be important in my future career. (R)</i>	.57	.58
<i>My English skills will help me get a better education (after school).</i>	.72	.72
<i>English will not play a big role in my further education (after school). (R)</i>	.59	.60
<i>With good English skills, I will be able to earn more money in future.</i>	.44	.44
<i>I'm only learning English because I don't want to get bad marks.</i>	.58	.61
<i>I would want to learn English, even if others (e.g. my parents or friends didn't think it was important.</i>	.48	.50
<i>I would want to learn English, even if it wasn't a mandatory subject. (R)</i>	.59	.60
<i>I only want to learn enough English to pass my exams.</i>	.63	.76
<i>I don't need English to pursue my personal interests. (R)</i>	.45	.46
<i>My English skills help me to learn more things independently.</i>	.39	.52
<i>I want to know English because it helps me pursue my interest.</i>	.80	.84
<i>I want to know English so that I can acquire new ideas and broaden my horizon.</i>	.62	.72
<i>I want to know English because, in the future, I would like to live abroad for some time (e.g. studying, working).</i>	.61	.61
<i>I don't need to know English well, because I don't enjoy visiting other countries. (R)</i>	.38	.50
<i>I want to know English because I want to travel to other countries.</i>	.78	.78
<i>I don't need to know English well, because I don't plan on ever living abroad. (R)</i>	.60	.65
<i>I want to know English because it helps you to meet people from other countries.</i>	.59	.69
<i>I want to know English so that I can exchange ideas with people from other countries.</i>	.65	.68
<i>I want to know English so that I can learn more about the culture and life of people from other countries.</i>	.67	.71

Table D.5: Language learning orientations Exploratory Factor Analysis. Factor loadings after rotation

No.	Item	Factor 1 Instrumental Obligation	Factor 2 International /Travel Orientation	Factor 3 Instrumental Benefit	Factor 4 Knowledge Orientation	Factor 5 Intrinsic Motivation
1.	<i>I only want to learn enough English to pass my exams.</i>	-.76	.01	-.15	-.14	.04
2.	<i>I don't really like English. (R)</i>	.69	-.17	.07	.01	.38
3.	<i>I'm only learning English because I don't want to get bad marks.</i>	-.65	-.05	-.02	.01	-.19
4.	<i>I don't need to know English well, because I don't plan on ever living abroad. (R)</i>	.47	.36	.34	-.06	-.16
5.	<i>I would want to learn English, even if it wasn't a mandatory subject. (R)</i>	-.41	-.25	.01	-.16	-.22
6.	<i>I would want to learn English, even if others (e.g. my parents or friends didn't think it was important.</i>	-.31	-.20	.06	-.16	-.29
7.	<i>I want to know English because I want to travel to other countries.</i>	-.05	.93	-.07	.05	-.10
8.	<i>I want to know English because it helps you to meet people from other countries.</i>	-.01	.67	.10	.07	.06
9.	<i>I want to know English so that I can learn more about the culture and life of people from other countries.</i>	.24	.66	-.11	.24	-.02
10.	<i>I don't need to know English well, because I don't enjoy visiting other countries. (R)</i>	-.03	.57	.12	-.14	.13
11.	<i>I want to know English so that I can exchange ideas with people from other countries.</i>	.05	.47	.08	.24	.25
12.	<i>English will not play a big role in my further education (after school). (R)</i>	.08	-.08	.75	.05	.02
13.	<i>I don't think that English will be important in my future career. (R)</i>	.12	.21	.62	-.07	.02
14.	<i>My English skills will help me get a better education (after school).</i>	.06	.10	.57	.29	.09

No.	Item	Factor 1 Instrumental Obligation	Factor 2 International Orientation	Factor 3 Instrumental Benefit	Factor 4 Knowledge Orientation	Factor 5 Intrinsic Motivation
15.	<i>I want to know English because, in the future, I would like to live abroad for some time ~~(e.g. studying, working)~~.</i> ³	.29	.10	.49	.12	.01
16.	<i>I think that knowing English will help me advance in my future career.</i>	- .15	.23	.48	.01	.43
17.	<i>With good English skills, I will be able to earn more money in future.</i>	- .04	- .03	.47	.35	.00
18.	<i>I want to know English because it helps me pursue my interest.</i>	.01	- .11	.19	.85	.02
19.	<i>I want to know English so that I can acquire new ideas and broaden my horizon.</i>	- .05	.13	.00	.70	.12
20.	<i>My English skills help me to learn more things independently.</i>	.12	.08	- .08	.54	.07
21.	<i>I don't need English to pursue my personal interests. (R)</i>	.14	.10	.24	.43	- .03
22.	<i>I find English exciting.</i>	.16	.00	.08	- .01	.79
23.	<i>I find learning English interesting.</i>	.30	.06	- .10	.15	.67
24.	<i>I think learning English is really fun and I see it as a challenge.</i>	.05	.04	.05	.31	.60

The items highlighted in grey are those that were included in the final questionnaire (Phase III). The numbers in bold indicate which factors these items most accurately represent.

³ The strikethrough text was excluded in the final questionnaire (Phase III) in an attempt to reduce the cross-loading onto the first factor, instrumental orientation.

D.3.3 Second Language Motivational Selves

Table D.6: Second language motivational selves Exploratory Factor Analysis. Factor loadings after rotation

No.	Item	Factor 1 Current L2 Self	Factor 2 Future L2 Self
1. Current	<i>I speak English fluently.</i>	.85	.08
1. Future	<i>I speak English fluently.</i>	.10	.73
2. Current	<i>I can communicate effortlessly with others in English.</i>	.90	-.07
2. Future	<i>I can communicate effortlessly with others in English.</i>	-.09	.91
3. Current	<i>I can express my whole personality in English.</i>	.84	.07
3. Future	<i>I can express my whole personality in English.</i>	.11	.78
4. Current	<i>I can write effortlessly in English.</i>	.83	.04
4. Future	<i>I can write effortlessly in English.</i>	-.02	.89
5. Current	<i>I can communicate successfully in English.</i>	.88	-.07
5. Future	<i>I can communicate successfully in English.</i>	-.19	.85
6. Current	<i>It's not a problem for me to use English.</i>	.89	.02
6. Future	<i>It's not a problem for me to use English.</i>	-.02	.92
7. Current	<i>I can read fluently in English.</i>	.83	-.09
7. Future	<i>I can read fluently in English.</i>	-.04	.89
8. Current	<i>I feel comfortable using English in different situations.</i>	.92	-.05
8. Future	<i>I feel comfortable using English in different situations.</i>	-.05	.89
9. Current	<i>Speaking English is part of who I am.</i>	.74	.12
9. Future	<i>Speaking English is part of who I am.</i>	.18	.73
10. Current	<i>I can express my thoughts and ideas without problems using English.</i>	.93	.00
10. Future	<i>I can express my thoughts and ideas without problems using English.</i>	.09	.78

The highlighted items are those that were included in the final questionnaire (Phase III). The numbers in bold indicate which factors these items most accurately represent.

D.3.4 Attitudes Towards Informal and Formal Language Learning

Table D.7: Attitudes towards informal and formal language learning Exploratory Factor Analysis. Factor loadings after rotation

No.	Item	Factor 1 Attitudes towards informal learning	Factor 2 Attitudes towards formal learning
1.	Most of my English knowledge comes from my free time.	-.78	-.08
2.	What I do in my free time does not really help me to learn English. (R)	-.75	.09
3.	The English I encounter in my free time is more useful to me than what we learn in school.	-.71	-.25
4.	Most of my English knowledge comes from school classes.	.67	.36
5.	I find the English that we learn in school more useful than what I encounter in my free time.	.54	.44
6.	I find the English I use outside of school less important than what we learn in school. (R)	-.49	-.19
7.	I think that the internet can help you improve your English.	-.46	.22
8.	In general, I think that you effectively learn English in school.	.06	.76
9.	What we do in English class really helps you to learn the language.	.12	.77
10.	When learning languages, it's important to know the grammar rules.	-.14	.77
11.	To me, English classes in school are a waste of time. (R)	.04	.56
12.	When learning languages, it's important to explicitly practice vocabulary.	.06	.43

The highlighted items are those that were included in the final questionnaire (Phase III). The numbers in bold indicate which factors these items most accurately represent.

D.3.5 Engagement

Table D.8: Affective, cognitive, and linguistic engagement Exploratory Factor Analysis. Factor loadings after rotation

Item	Factor 1 Cognitive Engagement	Factor 2 Linguistic Engagement	Factor 3 Affective Engagement
<i>All in all, I was [not at all – completely] focused on the content.</i>	.84	-.16	.07
<i>I was [not at all – completely] focused on trying to understand every single word.</i>	.84	.18	-.18
<i>The content was [not at all interesting – very interesting] to me.</i>	.73	-.03	.29
<i>During this activity, I [did not concentrate at all – concentrated hard] on it.</i>	.59	.20	.02
<i>All in all, I thought [very little – a lot] about the language during this activity.</i>	-.01	.89	.24
<i>All in all, I was [not at all – completely] focused on the language.</i>	.28	.55	-.11
<i>All in all, this activity was [boring – entertaining].</i>	-.18	.09	.86
<i>I [really did not like – really liked] this activity.</i>	.07	-.07	.84
<i>I felt [not at all – completely] absorbed in this activity.</i>	.32	.13	.49
<i>During this activity, I felt [confident – overwhelmed].</i>	.33	-.39	.46
<i>This activity was [very – not at all] beneficial to my English learning.</i>	.30	.12	.42

The highlighted items are those that were included in the final questionnaire (Phase III). The numbers in bold indicate which factors these items most accurately represent.

Appendix E. Findings Phase III

E.1 Preliminary Analyses: Normality and Outliers

Table E.1: C-tests descriptive statistics. Percentile scores by text, version, and time of data collection

Item	n		Missing		min		max		M		SD		S		K	
	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
VERSION 1																
Text 1	170	108	15	41	12.00	16.00	100.00	100.00	63.13	72.35	19.28	17.47	-0.46	-0.58	-0.11	0.18
Text 2	170	108	15	41	0.00	8.00	96.00	100.00	59.48	69.50	22.02	20.98	-0.32	-0.76	-0.59	-0.25
Text 3	170	108	15	41	0.00	0.00	98.00	100.00	55.21	66.98	22.35	22.00	-0.23	-0.37	-0.30	-0.26
Text 4	170	108	15	41	0.00	0.00	96.00	100.00	32.33	48.17	21.33	27.30	0.92	0.13	0.37	-1.03
Text 5	170	108	15	41	0.00	8.00	100.00	100.00	42.41	56.17	24.02	24.10	0.39	-0.07	-0.57	-1.02
Text 6	170	108	15	41	0.00	8.00	90.00	100.00	40.29	51.85	19.86	21.88	0.10	0.16	-0.04	-0.77
Text 7	170	108	15	41	0.00	0.00	96.00	100.00	36.24	57.37	24.28	21.50	0.35	-0.04	-0.50	-0.50
Total	170	108	15	41	8.86	25.43	90.00	98.00	47.01	60.34	17.01	18.25	0.57	0.12	-0.05	-0.73
VERSION 2																
Text 1	148	150	21	55	8.00	0.00	100.00	100.00	67.64	62.39	21.48	24.04	-0.48	-0.43	-0.31	-0.34
Text 2	148	150	21	55	22.00	20.00	100.00	100.00	79.72	77.39	15.71	16.95	-0.89	-1.31	0.45	2.45
Text 3	148	150	21	55	10.00	8.00	100.00	100.00	74.51	70.67	22.22	24.90	-0.74	-0.73	-0.45	-0.29
Text 4	148	150	21	55	0.00	0.00	100.00	100.00	56.97	53.03	25.76	28.45	0.08	-0.16	-1.12	-0.71
Text 5	148	150	21	55	0.00	0.00	96.00	96.00	36.01	42.81	30.89	28.63	0.16	0.39	-1.05	-1.20
Text 6	148	150	21	55	0.00	0.00	100.00	98.00	43.51	48.11	21.74	19.64	0.00	0.04	-0.01	-0.01
Text 7	148	150	21	55	0.00	0.00	100.00	100.00	39.05	54.45	33.01	28.15	-0.36	0.16	-0.87	-1.41
Total	148	150	21	55	23.71	6.86	97.43	96.57	56.77	58.60	17.27		-0.08	0.53	-0.61	-0.40

Table E.2: *Affective, cognitive, and linguistic engagement descriptive statistics*

<i>Item</i>	<i>n</i>	<i>Missing</i>	<i>min</i>	<i>max</i>	<i>M</i>	<i>SD</i>	<i>S</i>	<i>K</i>
<i>Affect 1: All in all, this activity was [boring — entertaining].</i>	260	94	40.60	100	82.80	12.26	-1.25	2.10
<i>Affect 2: The activity was [not at all interesting — very interesting] to me.</i>	260	94	22.06	100	75.75	15.64	-0.91	0.85
<i>Affect 3: I [really did not like — really liked] this activity.</i>	259	95	39.46	100	81.73	12.43	-1.01	1.24
<i>Cognition 1: During this activity, I [did not concentrate — concentrated hard] on it.</i>	261	93	8.57	100	62.76	17.99	-0.48	-0.66
<i>Cognition 2: I was [not at all — completely] focused on trying to understand every single word.</i>	261	93	0.00	100	49.18	20.21	0.07	-1.03
<i>Cognition 3: All in all, I was [not at all — completely] focused on the content.</i>	260	94	13.96	100	70.45	17.00	-0.95	0.08
<i>Language 1: All in all, I was [not at all — completely] focused on the language.</i>	261	93	3.71	100	48.70	19.18	0.02	-0.90
<i>Language 2: All in all, I thought [very little — a lot] about the language during this activity.</i>	257	97	0.00	94	39.49	20.02	0.33	-0.78
<i>Language 3: I was more concentrated on [understanding the overall meaning — understanding every single word].</i>	255	99	23.00	100	74.33	15.75	-1.07	0.99

Table E.3: *Behavioural engagement descriptive statistics*

<i>Variable</i>	<i>n</i>	<i>Missing</i>	<i>min</i>	<i>max</i>	<i>M</i>	<i>SD</i>	<i>S</i>	<i>K</i>
<i>Mean time spent on informal language practices per day (in minutes)</i>	262	92	7.60	360.00	83.05	58.93	1.80	4.86
<i>Mean number of unique informal language practices per day</i>	262	92	.14	4.00	1.29	.54	1.26	2.31

<i>1. I don't need English to pursue my personal interests. (R)</i>	335	287	19	67	7.53	6.60	2.63	2.79	-0.93	-0.51	-0.28	-1.00
<i>2. My English skills help me to learn more things independently.</i>	332	284	22	70	6.03	5.73	2.51	2.51	-0.10	-0.12	-0.94	-0.85
<i>3. I want to know English because it helps me pursue my interests.</i>	335	283	19	71	6.90	6.52	2.48	2.64	-0.55	-0.44	-0.66	-0.78
<i>4. I want to know English so that I can acquire new ideas and broaden my horizon.</i>	334	282	20	72	6.55	6.45	2.52	2.56	-0.46	-0.39	-0.66	-0.83
INTERNATIONAL ORIENTATION												
<i>1. I want to know English because I want to live abroad some time.</i>	335	287	19	67	5.76	5.54	2.93	3.03	-0.18	-0.08	-1.20	-1.21
<i>2. I want to know English because it helps you meet people from other countries.</i>	336	287	18	67	7.62	7.48	2.41	2.45	-1.00	-1.00	0.18	0.28
<i>3. I don't need to know English well, because I don't like traveling to other countries. (R)</i>	337	286	17	68	8.99	8.83	1.82	1.96	-2.21	-2.05	4.29	3.70
<i>4. I want to know English so that I can learn more about the culture and life of people from other countries.</i>	335	285	19	69	5.98	5.64	2.61	2.85	-0.23	-0.10	-0.91	-1.19
<i>5. I want to know English because I want to travel to other countries.</i>	336	285	18	69	8.16	7.65	2.20	2.57	-1.39	-1.07	1.43	0.16
<i>6. I don't need to know English well, because I don't ever plan on living abroad. (R)</i>	333	283	21	71	8.05	7.84	2.40	2.44	-1.43	-1.15	1.35	0.49

Note: Minimum and maximum values are not supplied in this table, as scores covered the whole range of possible values (1 through 10).

Table E.5: Motivational second language selves descriptive statistics

Item	n		Missing		Min		Max		M		SD		Skewness		Kurtosis	
	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
IDEAL FUTURE SELF																
1. I can communicate effortlessly with others in English.	327	285	27	69	5	4	10	10	9.22	9.10	1.22	1.31	-1.75	-1.62	2.78	2.33
2. I can express my whole personality in English.	327	285	27	69	3	2	10	10	8.62	8.38	1.69	1.84	-1.43	-1.31	1.82	1.47
3. I can communicate effectively in English.	330	285	24	69	5	4	10	10	9.35	9.21	1.16	1.30	-2.12	-1.80	4.25	2.69
4. Using English is not a problem for me.	329	285	25	69	4	4	10	10	9.24	9.09	1.31	1.46	-2.01	-1.78	3.71	2.75
5. I feel comfortable using English in different situations in my life.	330	285	24	69	3	2	10	10	8.90	8.56	1.62	1.97	-1.71	-1.53	2.49	1.85
6. I am able to express my thoughts and ideas effortlessly in English.	330	285	24	69	4	3	10	10	9.18	8.81	1.33	1.73	-1.80	-1.68	2.79	2.40
CURRENT SELF																
1. I can communicate effortlessly with others in English.	326	285	28	69	2	1	10	10	6.59	6.66	1.95	1.91	-0.31	-0.65	-0.65	0.12
2. I can express my whole personality in English.	327	286	27	68	1	1	10	10	6.08	5.99	2.08	2.14	-0.12	-0.20	-0.58	-0.50
3. I can communicate effectively in English.	330	285	24	69	1	1	10	10	7.15	7.20	2.05	2.00	-0.59	-0.73	-0.26	0.10
4. Using English is not a problem for me.	330	285	24	69	1	1	10	10	7.16	7.18	2.22	2.28	-0.57	-0.63	-0.55	-0.43
5. I feel comfortable using English in different situations in my life.	332	286	22	68	1	1	10	10	6.63	6.53	2.37	2.40	-0.35	-0.28	-0.70	-0.78
6. I am able to express my thoughts and ideas effortlessly in English.	331	285	23	69	1	1	10	10	6.31	6.16	2.10	2.14	-0.18	-0.29	-0.66	-0.64
DISTANCE (FUTURE – IDEAL SELF)																
1. I can communicate effortlessly with others in English.	325	285	29	69	-3	-5	7	9	2.66	2.45	1.93	1.86	0.12	0.50	-0.29	1.88
2. I can express my whole personality in English.	326	285	28	69	-7	-6	8	9	2.55	2.40	2.18	2.13	-0.17	0.09	1.21	1.32

<i>3. I can communicate effectively in English.</i>	328	285	26	69	-4	-6	9	9	2.21	2.01	2.06	2.07	0.48	0.33	0.28	1.80
<i>4. Using English is not a problem for me.</i>	328	285	26	69	-5	-6	8	9	2.09	1.91	2.12	2.29	0.38	0.37	0.25	1.47
<i>5. I feel comfortable using English in different situations in my life.</i>	329	285	25	69	-5	-8	9	9	2.29	2.02	2.32	2.48	0.33	0.00	0.21	1.46
<i>6. I am able to express my thoughts and ideas effortlessly in English.</i>	329	285	25	69	-6	-6	8	9	2.88	2.65	2.14	2.29	0.04	-0.27	0.13	1.85

Table E.6: Attitudes towards informal and formal learning descriptive statistics

Item	n		Missing		Min		Max		M		SD		Skewness		Kurtosis	
	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
ATTITUDES TOWARDS INFORMAL LEARNING																
1. I think that the Internet can help you improve your English.	326	287	28	67	5	4	10	10	8.97	8.87	1.24	1.39	-1.27	-1.21	1.39	0.90
2. The English I encounter in my free time is more useful than what I learn in school.	327	287	27	67	1	1	10	10	7.44	7.66	2.43	2.20	-0.78	-0.83	-0.27	0.01
3. The majority of my English skills come from my free time.	327	286	27	68	1	1	10	10	6.43	6.41	2.52	2.62	-0.13	-0.16	-1.09	-1.16
4. What I do in my free time doesn't really help me learn English. (R)	325	284	29	70	1	1	10	10	7.94	7.99	2.42	2.28	-1.14	-1.15	0.29	0.41
5. English, the way I use it outside of school, is less important than what we learn in class. (R)	322	282	32	72	1	1	10	10	7.54	7.54	2.13	2.13	-0.80	-0.96	0.05	0.51
ATTITUDES TOWARDS FORMAL LEARNING																
1. What we do in English class really helps you learn the language.	327	287	27	67	1	1	10	10	6.86	6.22	2.14	2.30	-0.69	-0.34	0.13	-0.56
2. When learning a language, it's important to know the grammar rules.	326	286	28	68	1	1	10	10	7.05	6.55	2.15	2.31	-0.54	-0.38	-0.20	-0.54
3. English classes in school are a waste of time. (R)	327	285	27	69	1	1	10	10	7.82	7.20	1.87	2.17	-1.08	-0.80	1.13	0.17
4. When learning a language it's important to practice vocabulary explicitly.	326	284	28	70	1	1	10	10	6.87	6.28	2.36	2.61	-0.58	-0.41	-0.49	-0.85
5. In general, I believe you learn English effectively in school.	324	284	30	70	1	1	10	10	6.31	5.80	1.97	2.21	-0.25	-0.19	-0.39	-0.69

E.2 Effects of Classroom Clustering

Intra-class correlations were inspected as a way of estimating effects of classroom clustering on the participants' C-test scores. In the table below, ICC1 denotes the amount of variance in individual item scores that can be explained by group membership. A common rule of thumb is to adjust for data clustering when this score exceeds .10 (Bliese, 2000). ICC2 is a measure of group-mean reliability, that is the probability with which groups can be reliably differentiated by their average scores. Correcting for data clustering is recommended where this probability is above .70 (ibid.).

<i>Variable</i>	<i>ICC1</i>	<i>ICC2</i>
C-TEST VERISON 1		
Text 1	.00	.00
Text 2	.23	.88
Text 3	.16	.82
Text 4	.24	.88
Text 5	.02	.35
Text 6	.03	.44
Text 7	.04	.46
C-TEST VERISON 2		
Text 1	.08	.67
Text 2	.04	.49
Text 3	.05	.57
Text 4	.06	.60
Text 5	.15	.81
Text 6	.09	.70
Text 7	.09	.71

E.3 Confirmatory Factor Analysis: C-tests

Figure E.1: C-test Confirmatory Factor Analysis, V1T1

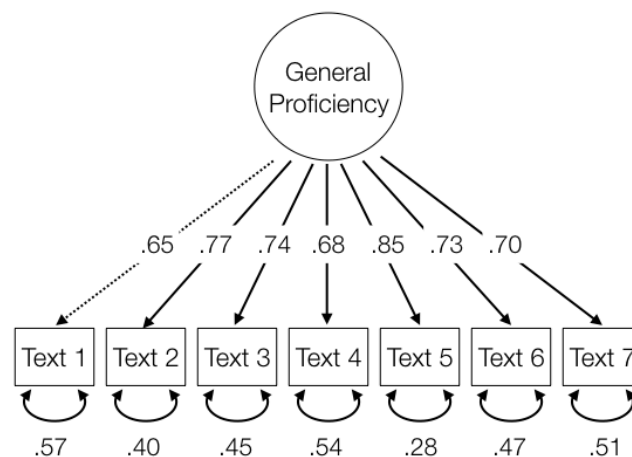


Figure E.2: C-test Confirmatory Factor Analysis, V1T1, including a correlation between the error variances of Text6 and 7

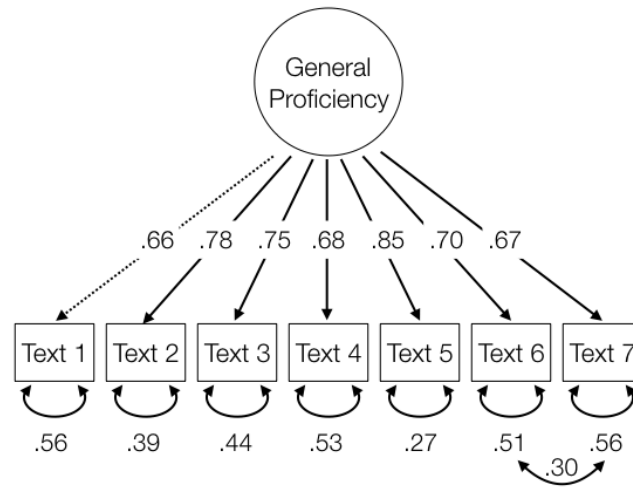


Figure E.3: C-test Confirmatory Factor Analysis, V2T1

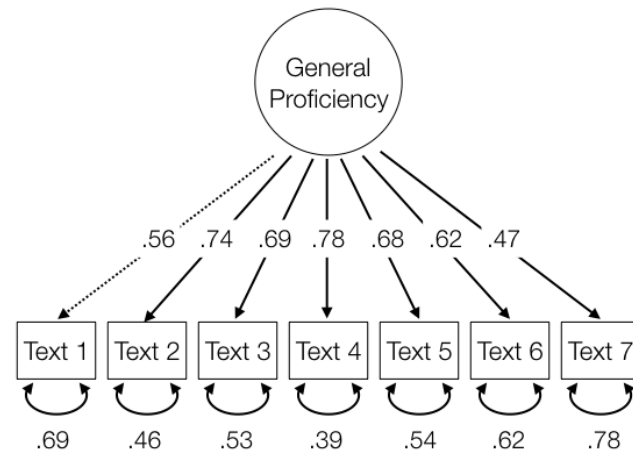


Figure E.4: C-test Confirmatory Factor Analysis, V2T1, including a correlation between the error variances of Text6 and 7

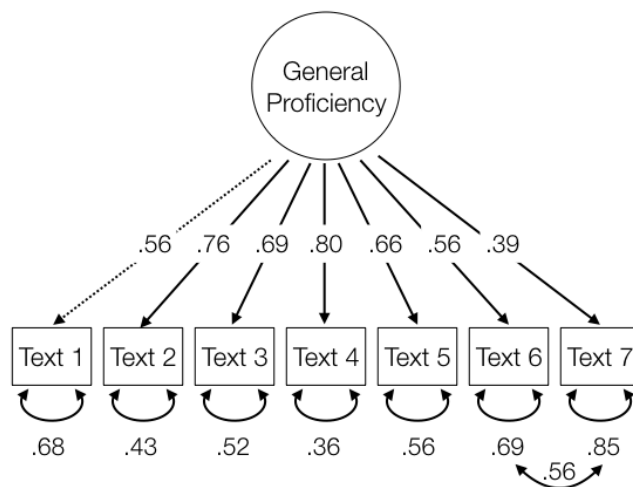


Figure E.5: C-test Confirmatory Factor Analysis, V1T2

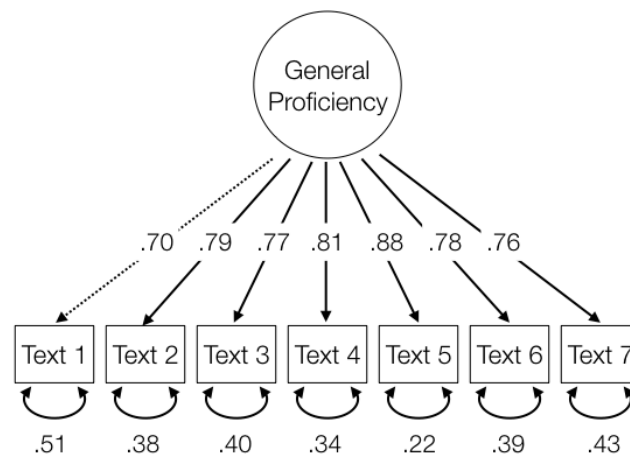
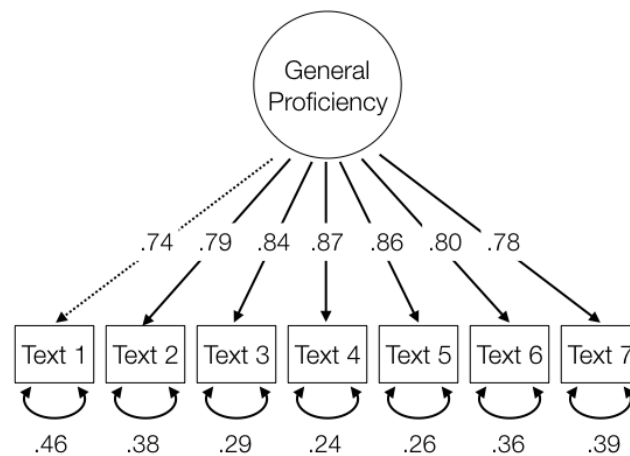


Figure E.6: C-test Confirmatory Factor Analysis, V2T2



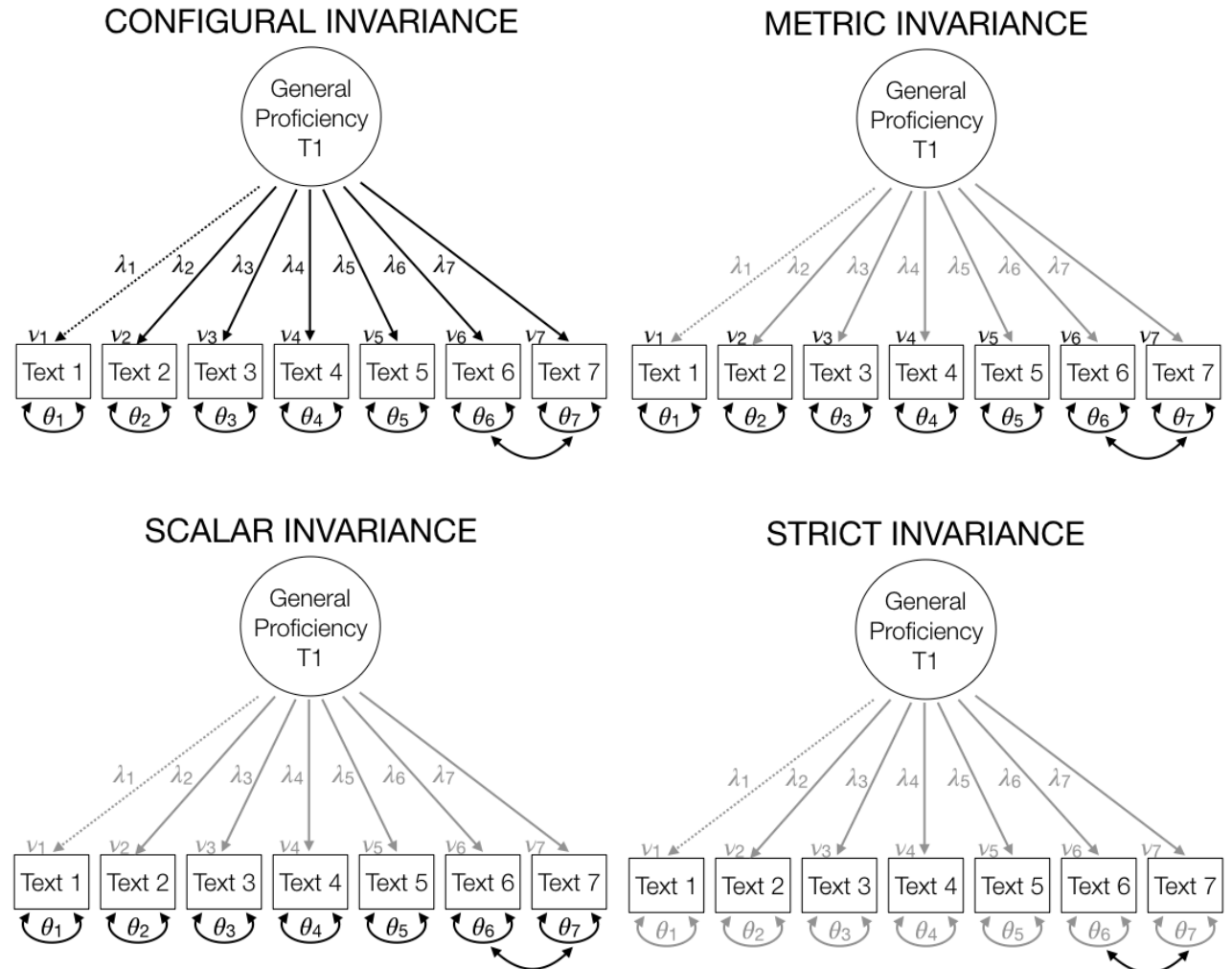
E.4 Measurement Invariance Modelling: C-tests

In the first step of MIM, the same factor structure is applied to both data sets as a baseline, to establish whether or not the same items measure the same latent constructs in both contexts (*configural invariance*). Next, *metric invariance* is established by constraining the factor loadings in both models to be equal, demonstrating that the items have the same meaning with regard to the latent construct in both instances. A good model fit at this stage suggests that analyses based on variances and covariances can be conducted, for example through regression modelling or SEM (Bialosiewicz, Murphy, & Berry, 2013). This, therefore, is the minimum level of invariance that should be achieved in the present study.

In the third step, the item intercepts are also constrained to be equal, which tests whether there are systematic differences between the two groups' scores on the measured items and, by implication, the latent factors (*scalar invariance*). In the MIM literature, achieving scalar invariance is seen as necessary for conducting mean comparison tests across groups or time (van de Schoot et al., 2012). Lastly, the error variances of the latent factors and manifest indicators are restricted to test *strict invariance*. A good model fit at this stage demonstrates that the latent factors are

measured identically in both models, but this level of invariance is rarely achieved in practice (ibid.).

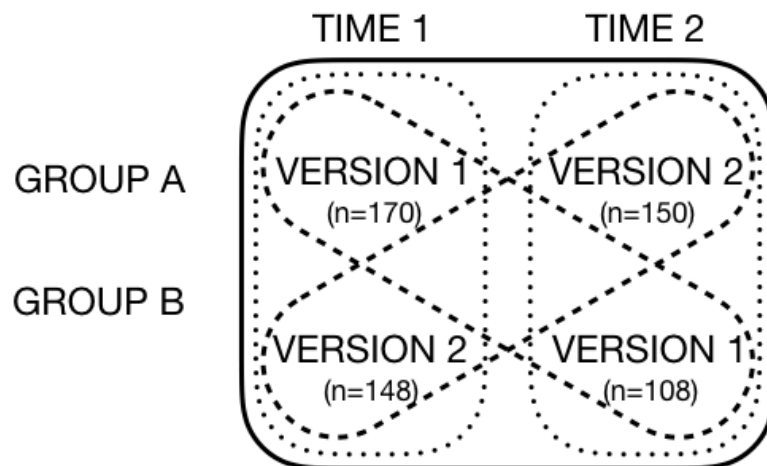
Figure E.7: Illustration of restrictions in Measurement Invariance Models



* Note: Greyed elements represent equality constraints across groups. λ represent the factor loadings, v the item intercepts, and θ the residual variances.

To account for the counter-balanced longitudinal design of this study, measurement invariance of the C-tests was tested in three ways (five analyses, see Figure D.8): invariance of the same versions across time and groups (dashed outlines), invariance between the two versions at time 1 and time 2 (dotted outlines); and finally, invariance across groups, time and different test versions (solid outline). The results of these analyses are discussed below and tables goodness-of-fit indices for all tests are presented, where the rows highlighted in grey mark the most restrictive models which achieved acceptable fit (CFI and TLI $\geq .90$; RMSEA and SRMR $\leq .10$).

Figure E.8: Group comparisons in C-test measurement invariance modelling



First, the invariance of both C-test versions across groups and time was tested (Figure D.8, dashed outlines). For both versions, the fit of the more restrictive metric invariance model was not significantly different from that of the configural invariance model, as indicated by the chi-square difference ($\Delta\chi^2_{V1}[6] = 11.43, p = .076$; $\Delta\chi^2_{V2}[6] = 4.25, p = .643$) and the excellent alternative fit indices at both levels. According to the chi-squared statistic, the fit of the scalar invariance model was significantly worse than the configural model for both versions ($\Delta\chi^2_{V1}[6] = 20, p < .01$; $\Delta\chi^2_{V2}[6] = 48.19, p < .001$). However, since the other fit indices still reached acceptable levels, this analysis supports the assumption that each of the C-tests had relatively good internal measurement validity, even when applied to different populations at different time points.

Table E.7: Goodness-of-fit indices, MIM of C-test versions across groups and time

Model	χ^2	df	p	CFI	TLI	RMSEA	SRMR
Version 1:							
$n_{T1} = 170, n_{T2} = 108$							
Configural Invariance	28.71	26	.425	1.00	1.01	.01	.02
Metric Invariance	40.14	32	.153	.99	.99	.04	.06
Scalar Invariance	63.93	38	.005	.98	.97	.07	.08
Version 2:							
$n_{T1} = 129, n_{T2} = 169$							
Configural Invariance	55.29	26	.001	.97	.96	.09	.04
Metric Invariance	59.54	32	.002	.98	.97	.08	.05
Scalar Invariance	107.73	38	< .001	.94	.93	.10	.08

Next, the model fit was compared across the two C-test versions (that is, across the two groups of participants) at T1 and T2 (Figure D.8, dotted outlines). While at T1, the chi-square statistic indicated a significantly worse performance as compared to the baseline ($\Delta\chi^2[6] = 19.26, p < .05$), the metric invariance model still achieved good fit in terms of the other indicators. The more restrictive scalar invariance model, however, did not achieve a good fit. These results indicate that the means of the two

groups cannot validly be compared at T1; however, it is not possible to determine from this data whether this is due to differences in the measurement quality of the two C-test versions or due to inherent differences between the participants in Groups A and B.

The researcher explored letting some intercepts vary freely across the two groups to try to achieve ‘partial scalar invariance’ (cf. Steenkamp & Baumgartner, 1998), as well as the effect of excluding some of the texts for which the two groups had substantively different mean scores. Scalar invariance could be achieved by excluding Texts 2–4 from the analysis altogether, but it was decided that this constituted too large a reduction in data, given that the planned SEM analyses could justifiably be conducted upon demonstrating only acceptable metric invariance (van de Schoot et al., 2012). The model comparing the two C-test versions at T2 similarly demonstrated good fit for metric invariance — although the chi-square difference between the metric and configural invariance model were once again significant ($\Delta\chi^2[6] = 20.90, p < .05$) — but also failed the test of scalar invariance.

Table E.8: Goodness-of-fit indices, MIM of C-test V1 and V2 at T1

<i>Model</i>	χ^2	<i>df</i>	<i>p</i>	<i>CFI</i>	<i>TLI</i>	<i>RMSEA</i>	<i>SRMR</i>
<i>Configural invariance</i>	30.01	26	.267	1.00	.99	.03	.03
<i>Metric invariance</i>	51.64	32	.015	.98	.97	.06	.07
<i>Scalar invariance</i>	214.74	38	< .001	.81	.79	.17	.15
INTERCEPT TEXT 4 FREED							
<i>Scalar invariance</i>	193.91	37	< .001	.83	.81	.16	.13
INTERCEPT TEXT 2 AND 4 FREED							
<i>Scalar invariance</i>	139.96	36	< .001	.89	.87	.14	.10
EXCLUDING TEXT 4							
<i>Configural invariance</i>	22.04	16	.142	.99	.99	.05	.03
<i>Metric invariance</i>	29.58	21	.101	.99	.98	.05	.05
<i>Scalar invariance</i>	169.41	26	< .001	.81	.78	.18	.13
EXCLUDING TEXT2 AND 4							
<i>Configural invariance</i>	9.09	8	.334	1.00	1.00	.03	.02
<i>Metric invariance</i>	12.43	12	.412	1.00	1.00	.02	.03
<i>Scalar invariance</i>	104.98	16	< .001	.84	.80	.19	.10
EXCLUDING TEXT 2, 3 AND 4							
<i>Configural invariance</i>	.065	2	.968	1.00	1.03	.00	.00
<i>Metric invariance</i>	1.19	5	.946	1.00	1.03	.00	.02
<i>Scalar invariance</i>	18.24	8	.020	.97	.95	.10	.05

Number of observations used: $n_{GrpA} = 170$, $n_{GrpB} = 148$.

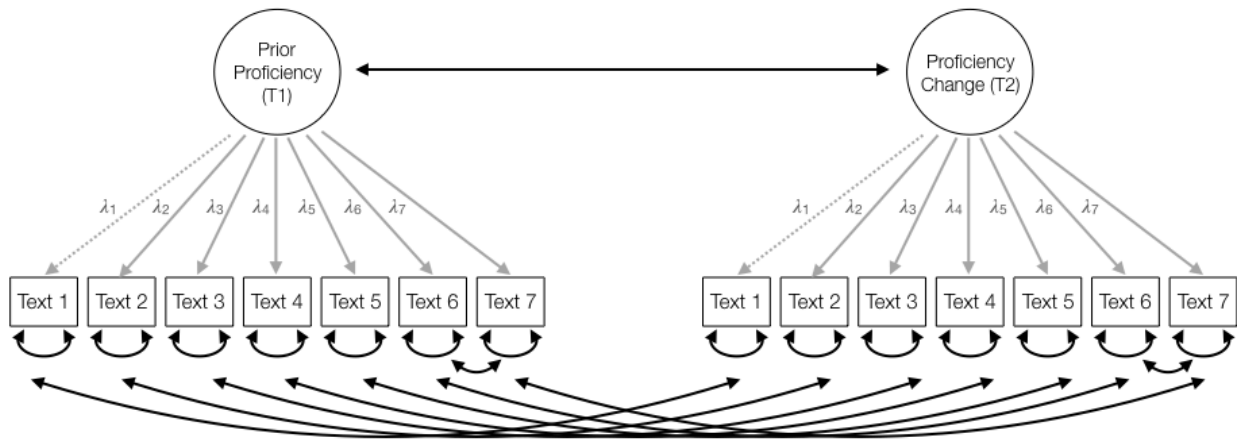
Table E.9: Goodness-of-fit indices, MIM of C-test V1 and V2 at T2

<i>Model</i>	χ^2	<i>df</i>	<i>p</i>	<i>CFI</i>	<i>TLI</i>	<i>RMSEA</i>	<i>SRMR</i>
<i>Configural invariance</i>	45.97	26	.009	.98	.97	.08	.03
<i>Metric invariance</i>	66.87	32	< .001	.97	.96	.09	.07
<i>Scalar invariance</i>	186.87	38	< .001	.88	.87	.17	.10

Number of observations used: $n_{GrpA} = 150$, $n_{GrpB} = 108$.

For the last analysis, which compared the fit of the one-latent factor model across time and groups (Figure D.8, solid outline), the tested model had to be slightly adapted to account for the longitudinal nature of the analysed data. A correlation between the latent factors representing proficiency at time 1 and time 2 was thus included, as well as residual correlations between the corresponding test items (Figure D.9).

Figure E.9: C-Test measurement invariance model: Within groups across time



This multi-group longitudinal invariance model achieved a good fit at the level of metric invariance, one that was not significantly different from the configural model fit ($\Delta\chi^2[6] = 10.53, p = .10$). Yet, the scalar invariance level did not achieve a good model fit, suggesting that a direct comparison of means across the two groups and time points would not be appropriate.

Table E.10: Goodness-of-fit Indices, MIM of C-test across time

Model	χ^2	df	p	CFI	TLI	RMSEA	SRMR
Configural Invariance	278.37	150	< .001	.94	.93	.08	.08
Metric Invariance	288.90	156	< .001	.94	.93	.08	.09
Scalar Invariance	510.14	173	< .001	.87	.86	.12	.14

Nevertheless, the fact that all MIMs achieved acceptable fit at the level of *metric invariance*, or beyond, strongly suggests that the two C-test versions provide valid data for the planned analyses based on variances and covariances, such as SEM. That is, based on the tests presented in this section, it seems justifiable to apply the same factor model to the data collected with both C-test versions, from both groups of participants, at both time points in order to derive a measure of general language proficiency.

E.5 Confirmatory Factor Analysis: Engagement

Figure E.10: Engagement Multi-level Confirmatory Factor Analysis. Amended two-factor model

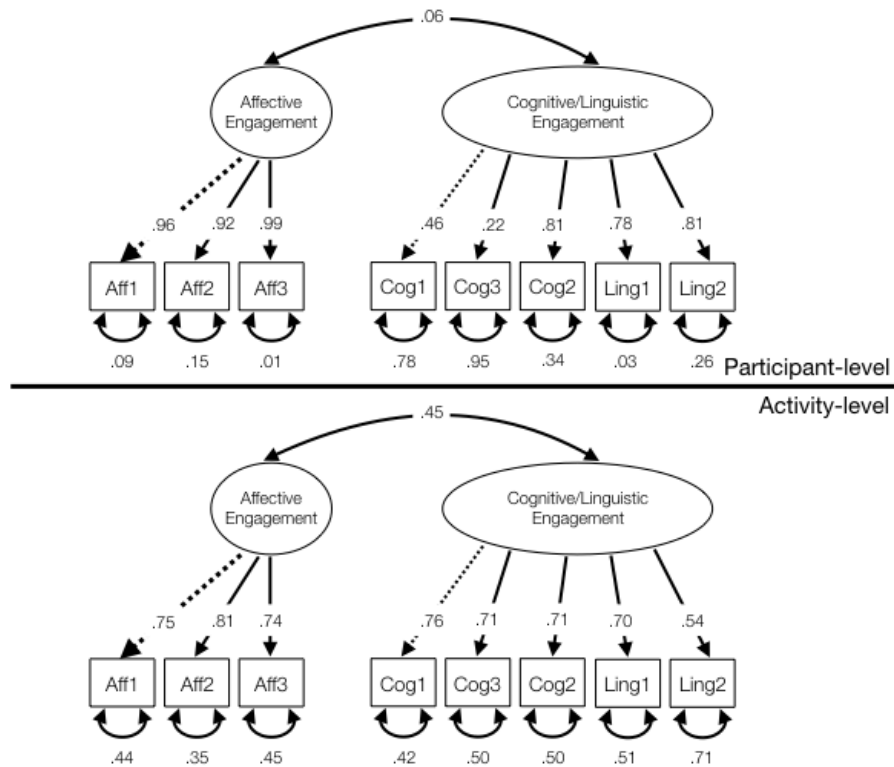


Figure E.11: Engagement Multi-level Confirmatory Factor Analysis (excluding data from participants who submitted entries for < 4 days)

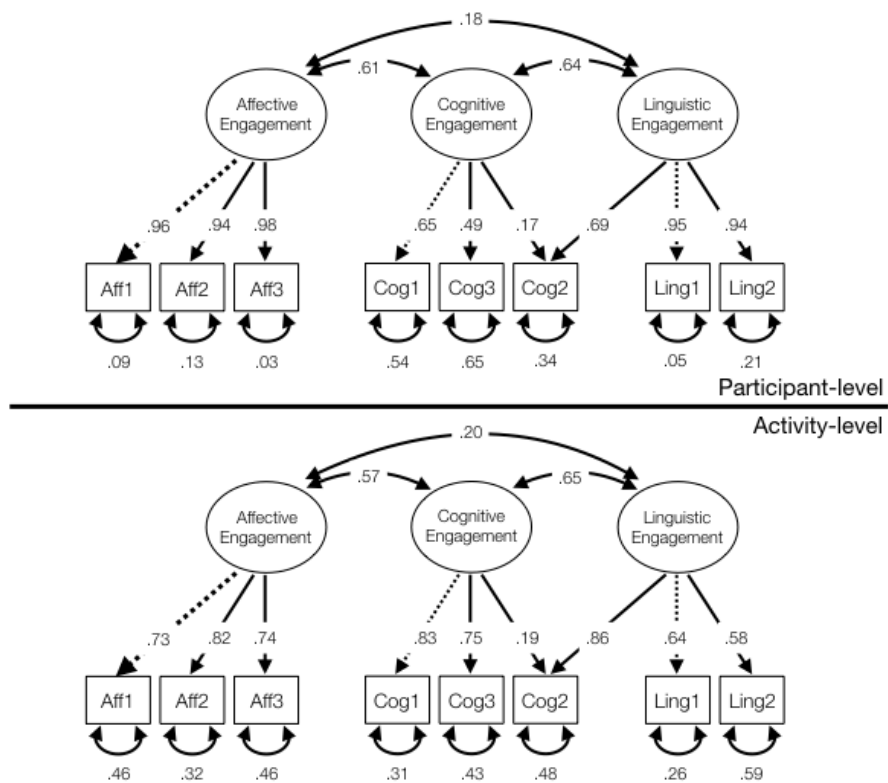


Figure E.12: Engagement Multi-level Confirmatory Factor Analysis (excluding data from participants who submitted entries for < 5 days)

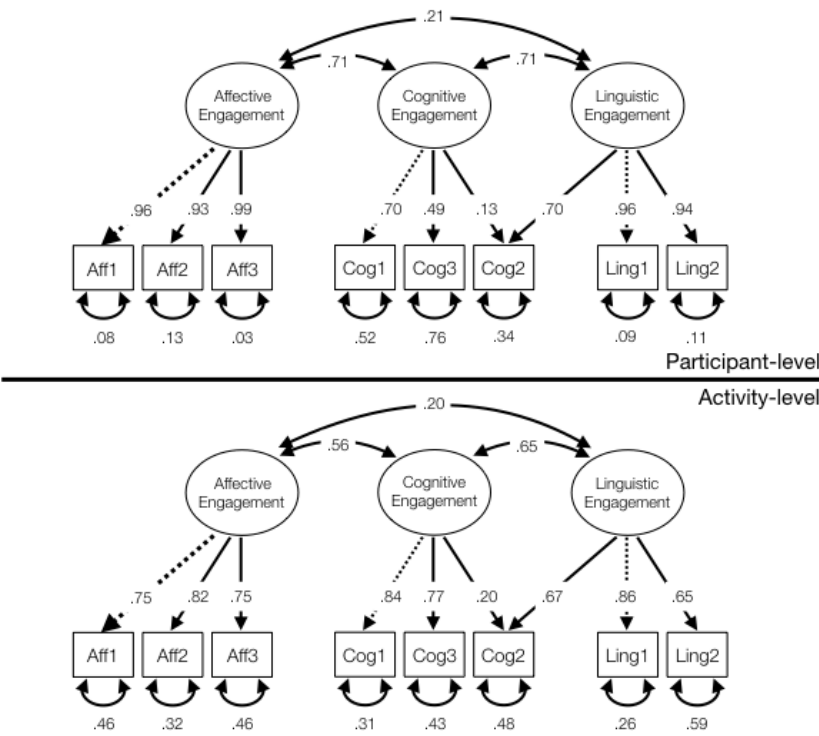


Table E.11: Engagement Multi-level Confirmatory Factor Analysis. Goodness-of-fit indices

Model	χ^2	df	p	$n_{w/i}$	$n_{b/w}$	CFI	TLI	RMSEA	$SRMR_{w/i}$	$SRMR_{b/w}$
All data	220.02	32	< .001	1505	254	.97	.95	.05	.04	.06
Excl. participants with entries over ≤ 3 days	96.97	32	< .001	1385	190	.98	.97	.05	.04	.05
Excl. participants with entries over ≤ 4 days	107.18	32	< .001	1274	163	.97	.96	.05	.04	.05

Appendix F. Findings Phase III, RQ 1: How do These Learners Engage in English Informal Second Language Practices?

F.1 Deriving Measures of Behavioural Engagement

Table F.1: Number and proportion of diary entries in each category of ISLPs

Activity	Total number of entries	Proportion of total entries
1. Listening to music (background)	218	12.10%
2. Listening to music (focused)	85	4.72%
3. Listening to music and reading lyrics	15	.83%
4. Listening to music and singing along	285	15.82%
5. Listening to music, reading lyrics, and singing along	43	2.39%
6. Video-based social media	262	14.54%
7. Text-based social media	77	4.27%
8. Mixed social media	256	14.21%
9. Films without subtitles	26	1.44%
10. Films with English subtitles	3	.17%
11. Films with German subtitles	6	.33%
12. TV programs without subtitles	97	5.38%
13. TV programs with English subtitles	43	2.39%
14. TV programs with German subtitles	15	.82%
15. Single-player video game	25	1.39%
16. Multi-player video game	19	1.05%
17. Multi-player video game and communicating via text chat	17	.94%
18. Multi-player video game and communicating via voice chat	16	.89%
19. Multi-player video game and communicating via text and voice chat	17	.94%
20. Having a direct conversation (in person)	52	2.89%
21. Having a remote conversation (e.g. telephone, voice chat)	32	1.78%
22. Reading books	27	1.50%
23. Reading magazine or news articles	11	.61%
24. Reading other online texts	33	1.83%
25. Other reading	9	.50%
26. Writing short texts or messages (e.g. WhatsApp, messenger)	45	2.50%
27. Writing longer texts	20	1.11%
28. Listening to a conversation	17	.94%
29. Listening to audiobooks	1	.06%
30. Listening to podcasts	8	.44%
31. Listening to radio programs	4	.22%
32. Other (computer programming)	2	.11%
Various	16	.89%
Total	1,802	

F.2 Popularity of Different Activities

Table F.2: Breakdown of ISLPs in terms of frequency, total duration, duration per session and prevalence within the sample

Activity	Frequency		Total duration		Duration per session	Prevalence
	Total number of entries	Proportion of entries in this category/ in total	Total time spent in minutes	Proportion of duration in this category/ total duration	Time per entry in minutes M (SD)	Proportion of students who engaged in the activity at least once
FILMS						
without subtitles	26	74.29%	2,831	81.77%	108.89 (58.14)	8.02%
with English captions	3	8.57%	329	9.50%	109.67 (20.01)	1.15%
with German subtitles	6	17.14%	302	8.72%	50.33 (45.17)	2.29%
Total: Films	35	1.93%	3,462	2.43%	98.91 (57.55)	10.42%
SERIALS & TV SHOWS						
without subtitles	97	62.58%	10,675	65.86%	113.56 (81.25)	16.79%
with English captions	43	27.74%	3,340	20.61%	77.67 (37.25)	8.02%
with German subtitles	15	9.68%	2,193	13.53%	156.64 (96.66)	3.05%
Total: Serials & TV	155	8.57%	16,208	11.38%	107.34 (76.12)	27.03%
SOCIAL MEDIA						
YouTube	320	52.46%	19,757	62.81%	62.72 (62.56)	51.14%
Instagram	225	36.89%	9,470	30.10%	43.24 (43.23)	41.60%
Twitter	22	3.61%	810	2.57%	36.82 (23.07)	3.62%
Snapchat	17	2.79%	435	1.38%	27.19 (17.51)	4.58%
Facebook	7	1.15%	200	.64%	33.33 (15.06)	1.15%
Other	19	3.11%	785	2.50%	71.36 (50.20)	3.05%

<i>Activity</i>	Frequency		Total duration		Duration per session	Prevalence
	<i>Total number of entries</i>	<i>Proportion of entries in this category/ in total</i>	<i>Total time spent in minutes</i>	<i>Proportion of duration in this category/ total duration</i>	<i>Time per entry in minutes M (SD)</i>	<i>Proportion of students who engaged in the activity at least once</i>
Type of social media						
<i>Video-based</i>	262	44.03%	15,606	49.99%	60.48 (63.49)	47.33%
<i>Text-based</i>	77	12.94%	2,373	7.60%	31.34 (27.91)	19.47%
<i>Mixed</i>	256	43.04%	13,238	42.41%	53.26 (49.13)	41.98%
Total: Social media	610	33.72%	31,587	22.17%	53.86 (55.51)	78.24%
WRITING						
<i>Shorter texts (e.g. brief notes and text messages)</i>	45	69.23%	2,036	64.27%	45.26 (78.20)	10.69%
<i>Longer texts</i>	20	30.77%	1,132	35.73%	61.56 (71.18)	6.49%
Total: Writing	65	3.59%	3,168	2.22%	51.93 (78.08)	14.44%
SPEAKING						
In person						
<i>with acquaintances</i>	50	96.15%	2,222	99.42%	44.43 (48.98)	13.74%
<i>with strangers</i>	2	3.85%	13	.58%	6.50 (4.95)	.76%
Technology-mediated						
<i>with acquaintances</i>	20	62.50%	2,649	67.44%	132.45 (127.79)	3.44%
<i>with strangers</i>	12	37.50%	1,279	32.56%	116.27 (87.07)	3.05%
Total: Speaking	84	4.64%	6,163	4.33%	74.25 (88.65)	19.69%

<i>Activity</i>	Frequency		Total duration		Duration per session	Prevalence
	<i>Total number of entries</i>	<i>Proportion of entries in this category/ in total</i>	<i>Total time spent in minutes</i>	<i>Proportion of duration in this category/ total duration</i>	<i>Time per entry in minutes M (SD)</i>	<i>Proportion of students who engaged in the activity at least once</i>
MUSIC						
<i>Listening to music (background)</i>	370	57.81%	36,990	64.26%	99.97 (90.58)	62.60%
<i>Listening to music (focused)</i>	270	42.19%	20,571	35.74%	76.19 (68.78)	56.11%
Total: Music	640	35.38%	57,561	40.40%	89.85 (82.84)	83.21%
VIDEO GAMES						
<i>Single-player games</i>	25	26.60%	3,160	23.51%	131.67 (98.71)	5.73%
<i>Multi-player games</i>	69	73.40%	10,280	76.49%	155.76 (103.82)	13.79%
Communication with others						
<i>Video games + text or voice chat</i>	50	55.56%	8,500	63.24%	123.50 (84.07)	11.20%
<i>Just gaming</i>	40	44.44%	4,940	36.76%	170.00 (111.69)	8.88%
Total: Video Games	94	5.20%	13,440	9.43%	149.33 (102.49)	18.30%
READING						
<i>Books</i>	27	34.18%	2,396	49.07%	88.74 (113.57)	6.73%
<i>Magazines or news articles (paper or online)</i>	11	13.92%	485	9.93%	44.09 (67.74)	3.05%
<i>Other online texts</i>	33	41.77%	1,604	32.85%	48.61 (63.80)	11.45%
<i>Other</i>	8	10.13%	398	8.15%	77.56 (103.17)	2.67%

Total: Reading	79	4.37%	4,883	3.43%	61.81 (85.55)	20.46%
<i>Activity</i>	Frequency		Total duration		Duration per session	Prevalence
	<i>Total number of entries</i>	<i>Proportion of entries in this category/ in total</i>	<i>Total time spent in minutes</i>	<i>Proportion of duration in this category/ total duration</i>	<i>Time per entry in minutes M (SD)</i>	<i>Proportion of students who engaged in the activity at least once</i>
LISTENING						
<i>Audiobooks</i>	1	3.33%	20	1.18%	—	.38%
<i>Radio programs</i>	4	13.33%	175	10.42%	43.75 (23.59)	1.15%
<i>Podcasts</i>	8	26.67%	325	19.35%	40.63 (35.90)	2.29%
<i>Overheard conversations</i>	17	56.67%	1,160	69.05%	72.50 (95.64)	4.20%
Total: Listening	30	1.66%	1,680	1.18%	55.29 (67.49)	7.03%
<i>Other</i>	1	.06%	60	.04%	—	.38%
<i>Various *</i>	16	.88%	4,250	2.98%	—	1.53%
ALL CODED ENTRIES	1,809		142,462		78.50 (78.86)	

Note: The 'total' numbers for frequency and duration may not always represent the sums of the preceding values, since some participants submitted diary entries without answering all of the follow-up questions. For example, a participant may have reported they were reading for 30 minutes, but not recorded what type of text they read. To maximise the available data, this entry would be counted towards the 'Total: Reading' scores without being in the preceding rows.

F.3 Findings Regarding Specific Informal Second Language Practices

F.3.1 Film and Television

As could be expected, based on the findings of previous studies and the Phase I interviews, the participants altogether spent substantially more time watching serials (11.4% of the total duration of entries) than films (2.43%). When asked how they accessed video content (excluding online videos discussed under *social media*), participants most frequently reported using internet streaming services (93.1% of all film/TV entries). Only two entries were made in which participants recorded watching English-language TV broadcasts. This was unsurprising, given that almost all programmes shown on German television are dubbed, with the exception of content broadcast on English-language channels, such as BBC World and Comedy Central, which are only available through premium TV packages. Less expected was that not a single participant reported watching videos downloaded from the internet, which was very common in previous studies of ISLL in France (Sokkett, 2014b). This discrepancy could indicate a cultural difference between Germany and France or a significant shift in viewing practices in recent years, given the wider availability of legal, paid online streaming services. However, considering how seriously media piracy is treated in German public discourse, it is also conceivable that the participants in this study simply did not feel comfortable reporting that they engaged in (illegal) downloading of content.

Besides the way in which learners access content, a common topic of research is the use of L2 captions and L1 subtitles. In this study, the data revealed that the vast majority of both films (81.8% of the relevant entries) and serials (65.9%) were watched without the use of any subtitles. Nevertheless, subtitles in both German and English were markedly more popular when watching serials than with films. Participants' remarks in the focus groups indicated that this difference might relate to the relative narrative complexity of the two media: Several pupils described using English-language captions to help them follow complex narratives in serial dramas, such as *Sherlock* or *Breaking Bad*, rather than to help them overcome language barriers per se.

Of the descriptive statistics concerning the use of captions and subtitles, the duration per session of the entries relating to films with German subtitles stands out particularly: At 50.3 minutes, the mean length of time in this category is substantively below the average running time of a movie (around 90 minutes: Olson, 2014). This could indicate that students may get frustrated with processing English speech and German text simultaneously — something for which many focus group participants expressed a strong dislike — and stop watching the film before the end. Looking at the individual data points, this seems likely: Two participants seemed to have watched a full movie (entries of 90–120 minutes), whereas four others did not (10–40 minutes). Meanwhile, almost three times as many entries were made for watching serials with English captions than with German subtitles (43 vs. 15), but the average viewing time in the latter category was noticeably higher ($M = 156.64$ vs. $M = 77.67$). This stark difference could be attributable to individual learners' viewing

habits: Nine out of 15 entries for watching serials with English subtitles were made by three learners, who seemed to tend towards longer viewing times ('binge-watching'), between one and four hours per session. Of course, the conclusions drawn in this paragraph should be treated only as speculation, particularly given that they are based on very small numbers of diary entries.

All in all, the duration of the recorded 'sessions' of watching films and serials, as well as the recorded levels of affective engagement, were among the highest in comparison to other types of activities, being significantly behind only playing video games while concurrently chatting with other players. Regarding the levels of cognitive and linguistic engagement, the emerging picture is less clear, with mid-range values being reported, which were not significantly different from most other activities.

F.3.2 Social Media

Little specific data about L2 social media use has been available from previous studies on ISLL, as many have subsumed social networking under the broader category of 'browsing the web' (e.g. Sockett, 2014a; Sundqvist, 2009). To get a better idea of the participants' social media use in this study, they were asked to record which platform they visited and the specific ways in which they used English there (Table F.3).

Overall, social media use was the second most popular activity in terms of frequency, duration, and prevalence behind listening to music. The vast majority of L2 social networking activities took place on YouTube (62.8% of the total time, 52.5% of all entries) and Instagram (36.9% of time, 30.1% of entries). Both of these platforms were used, at least once, by a large proportion of the participants (51.1% and 41.6% respectively). Other generally popular social media platforms, including Twitter and Facebook, however, were used (at least in English) by only a few individual students. It was also surprising to see only one participant mentioning the use of Facebook Messenger for synchronous, written communication in English. It seems likely that others would have also used text chat to stay in touch with international friend, but through services such as WhatsApp, which was coded not in the social media category, but as a 'short written exchange' (Section F.3.3 below). Finally, participants' open-ended answers indicated that other platforms they used included *9GAG* (5 entries), *Pinterest* (2 entries), *Tumblr*, *musical.ly*, and *Bottled* (1 entry each).

Table F.3 gives an overview of the specific ways in which the participants used their L2 on social media. It must be noted that participants were able to select more than one activity at a time, so some entries are counted in more than one column. Furthermore, if a participant submitted an entry in which they reported watching videos and reading other users' comments on Instagram for 45 minutes, this does not mean that they spent the full 45 minutes both listening to and reading English. That is, inferences cannot be drawn from this data about the *amount of time* participants spent using language in particular ways.

Watching videos was by far the most frequent activity across all major platforms. This was expected for YouTube, Instagram and Snapchat, as all of them are — to a greater or lesser extent — intended for video sharing. Even so, a large proportion (of the admittedly small number) of entries about Twitter and Facebook use also mentioned watching videos as one of the main activities. Meanwhile, only just over 50% of the social media entries mentioned any form of reading. This is surprising, given that this language skill is involved in a wide range of activities including reading status updates, video/image descriptions, user comments, video subtitles, etc. This data does seem to show that, even though the internet is commonly thought of as a text-based medium, the current ‘Web 2.0’ landscape is focused largely on visual and auditory media.

At its heart, this study is concerned with the ways in which learners utilise their L2 beyond the classroom. To reflect this, the researcher decided to recode the social media data into language-focused categories: Instead of distinguishing an activity by the app or website visited, social media use was categorised by whether it was video-based (watching videos was the *only* activity), text-based (reading and writing were the *only* activities), or mixed (involving both watching videos and reading/writing). In terms of the total duration, half of all social media use in this study involved participants *only* watching videos. Of the remaining time, 42.2% involved mixed social media use, meaning the students watched videos as well as read or wrote comments, status updates, etc. Only a very small part of social media use (7.6%) was purely text based and, given that it took place largely on Instagram, it was likely to include an abundance of (still) images in addition to the textual components.

All three forms of social media use — video- text-based and mixed — were among the activities on which the participants spent the lowest continuous amounts of time (‘per session’), with text-based social media significantly below the other two types. The affective engagement scores for entries in these categories not differ significantly from most other activities, although text-based social media entries also fell significantly below video-based social media entries in this regard. In terms of cognitive and linguistic engagement, these three categories were indistinguishable from most other activities, including predominantly mid-range scores.

Table F.3: Number and proportion of entries reporting specific activities engaged in on various social media platforms

Social media platform	Total number of entries for this platform	Number of entries with activities specified	Watching videos		Reading subtitles		Reading comments		Reading image headers or status updates		Reading longer texts (e.g. external links, Snapchat Discovery)		Writing comments or status updates		Synchronous written communication (text chat)	
YouTube	320	320	320	100.00%	29	9.06%	97	30.31%	—	—	—	—	17	5.31%	—	—
Instagram	225	224	170	75.89%	—	—	154	68.75%	97	43.30%	—	—	23	10.27%	—	—
Twitter	22	22	11	50.00%	—	—	—	—	21	95.45%	11	50.00%	3	13.63%	—	—
Snapchat	17	16	12	75.00%	—	—	6	37.50%	6	37.50%	5	31.25%	7	43.75%	—	—
Facebook	7	6	3	50.00%	—	—	0	0%	2	33.33%	0	0%	0	0%	1	16.67%
Other	19	11	2	18.18%	—	—	0	0%	9	81.81%	0	0%	0	0%	1	9.09%
Total	610	599	518	86.48%	29	4.84%	257	42.90%	135	22.54%	16	2.67%	50	8.35%	2	.33%
Total per language skill	Listening				Reading				Writing				Written communication			
	610	599	518	86.48%					326	54.42%			50	8.35%	2	.33%

Note: Participants were able to select more than one activity, so that some entries are counted in more than one activity column.

A dash (—) indicates a category that was not selected during any diary entries.

Table F.4: Number and proportion of entries reporting specific activities while playing video games

Game type	Total number of entries for this game type	Number of entries with activities specified	Listening to spoken English		Reading English texts		Writing English (e.g. text commands)		Communicating with other players in English through text chat		Communicating with other players in English through voice chat	
Single-player	25	25	14	56.00%	22	88.00%	11	44.00%	—	—	—	—
Multi-player	69	69	40	57.97%	45	65.22%	31	44.93%	34	49.28%	33	47.83%
Total	94	94	54	57.45%	67	71.28%	42	44.68%	34	36.17%	33	35.11%

Note: Participants were able to select more than one activity, so that some entries are counted in more than one activity column.

A dash (—) indicates a category that was not selected during any diary entries.

F.3.3 Writing

Writing was not a particularly popular activity in this study, accounting for only 65 diary entries (3.6%) and 3.168 minutes of the total time spent on ISLPs (2.2%). In fact, in terms of these criteria as well as the prevalence (14.4% of participants made at least one entry on writing), writing ranked second to last in popularity among all activities, with only listening being less popular in terms of frequency and duration of entries, and films in terms prevalence.

Overall, writing was among the activities with the lowest comparable duration 'per session'. Unsurprisingly, participants spent more time per session on writing longer texts such as e-mails, blog posts, or journal entries than on writing shorter texts, including short notes or WhatsApp messages. The latter category, on the other hand, was more popular in terms of the frequency, total duration, and prevalence than writing longer messages. It should be kept in mind, however, that particularly text messaging, even though it was coded as a writing activity in this study, is likely to also include a good amount of reading, for example of the interlocutor's messages.

The levels of affective engagement reported for writing activities were comparatively low — though significantly higher than both listening to spoken word and to background music. In terms of cognitive engagement, however, writing was among the activity with the highest scores and the reported levels of linguistic engagement were significantly higher than those for any other activity.

F.3.4 Speaking

Even though, in the focus group interviews, students frequently mentioned that they rarely or never spoke English outside of the classroom, speaking English was not an uncommon activity in this Phase of the study. It was approximately equal in popularity to video games in terms of the proportion of students who reported engaging in this activity at least once (19.5%) and the number of entries made (4.6% of the total, ranked fifth). However, speaking accounted for only half as much of the total time spent as video games (6,163 minutes/4.3% vs. 13,790 minutes/9.6%) since the average duration per entry was also half as long. On the whole, however, the time reported in each entry on speaking was comparable to most other activities.

The participants were asked to clarify whether they held the conversations in person or remotely, and how well they were acquainted with their interlocutors. This was to find out more about the role that technology played in giving the participants an opportunity to speak English. On the one hand, conversations in person were more frequent than through the mediation of technology (52 vs. 32 entries) and part of the ISL practices of a larger proportion of the students (14.1% vs. 6.5%). On the other hand, the time the participants spent speaking remotely to others far outweighed that spent in direct conversation (c. 65.5 hours vs. 37.3 hours), due to the significantly longer average durations per conversation (in minutes: $M = 126.71, SD = 113.72$ vs. $M = 42.97, SD = 48.58$).

The relative prevalence of direct conversations was surprising, given that in the focus groups the participants had said they only spoke English in person if they

happened to have friends, family members, or exchange students visiting from abroad. They also mentioned that there was very little opportunity to speak English for those who were not personally acquainted with English speakers, as there were very few foreign visitors or tourists in the small towns in which they lived. This was well reflected in the data presented above, given that the occurrence of in-person conversations with strangers was vanishingly rare (two entries). By contrast, about a third of the technology-mediated conversations were with people whom the participants had never met in person before, for example online friends. The data show that, while technology-mediated conversations might not be very prevalent within the sample as a whole, some students do seem to take advantage of the affordances that technology provides, with regard to allowing them to meet and speak to others in English.

On the whole, speaking did not differ from most other activities in terms of the affective engagement that participants felt, but it was among the group of activities with the highest levels of cognitive engagement and second only to writing in terms of participants' focus on language.

F.3.5 Music

In the current study, as in previous ISLL research, listening to music made up the by far largest proportion of the recorded time spent on ISLPs (40.4%). The time spent listening to music varied greatly between participants, ranging from 2 minutes up to 10 hours ($M = 89.95\text{min}$, $SD = 82.94$) with most records clustering in the 30–120 minute range (Figure F.1). However, only 35.5% of the time dedicated to music was comprised of 'focused listening', as classified by the participants themselves. Most of the time (64.5%), participants reported playing music 'in the background', while also doing something else. What these concurrent activities were and whether they also involved ISLL was not recorded. Unsurprisingly then, listening to music in the background scored significantly below all other activities — including focused listening — in terms of the affective, cognitive, and linguistic engagement felt by the participants who reported it. This finding lends support to the validity of the participants' self-reporting of an activity as either focused or unfocused listening.

Besides whether or not they focused on the music, participants were also asked if they had been singing along or reading the song lyrics (Table F.5). In fact, there were almost as many entries which mentioned singing along (44.1%) as entries reporting just listening to music (46.9% for focused and background listening combined). These data also show that participants were looking at song lyrics mostly while also singing (6.7%), not as an isolated activity (2.3%). As with the entries pertaining to social media, no inferences can be drawn from this data about the amount of time participants spent on each of these specific activities.

Figure F.1: Duration of entries pertaining to music listening

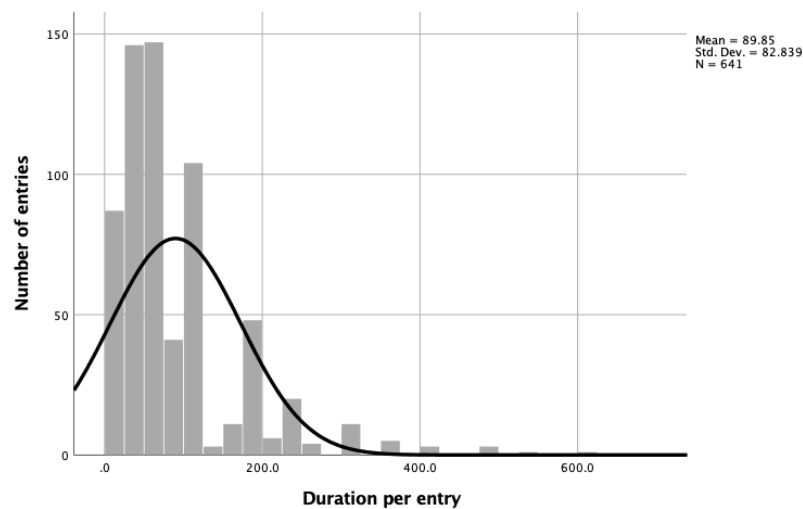


Table F.5: Number and proportion of entries reporting specific activities while listening to music

Activity	Number of Entries	Proportion
<i>Singing along to music</i>	285	44.12%
<i>Listening only (in the background)</i>	218	33.75%
<i>Listening only (focused)</i>	85	13.16%
<i>Reading lyrics and singing along</i>	43	6.66%
<i>Reading lyrics</i>	15	2.32%

F.3.6 Video Games

In terms of the total duration, video games seemed to be relatively popular in this study, ranking 4th among the activities (9.4% of all entries). Meanwhile, the other criteria revealed that games accounted for a disproportionately smaller number of the total entries (5.2%) and were actually played by very few participants (18.3% of the sample). The high ranking in terms of total duration can be attributed to the fact that this category had by far the highest average duration per session, albeit with large individual variations ($M = 148.04, SD = 101.83$). This pattern, a few active gamers spending large consecutive amounts of time on video gaming, replicates what has already been well documented in previous ISL literature (e.g. Chik, 2014; Sundqvist, 2009; Sundqvist & Wikström, 2015).

Participants recorded what type of game they played and which language skills they used during gameplay (Table F.4, page 461). In this study, multi-player games were much more popular than single-player games, accounting for 73.4% of all entries. Regarding language skills, the data show that reading — and to a lesser extent listening and writing — was an important part of gameplay. This is not surprising, given that all instructions and commands in English-language games are likely to be written in the L2. It also makes sense that a larger proportion of single-player than multi-player game entries (88.0% vs. 65.2%) were concerned with reading, because the former tend to be more narrative and include cutscenes (short, non-interactive video sequences, which introduce new gameplay elements or contribute to in-game

story-telling). Such sequences often contain English-language subtitles, which the participants would have read. For the same reason, one might also have expected listening to be a larger part of single-player gameplay, but this was not the case. One explanation could be that the students could have muted the sound or skipped video sequences in single-player games.

Of the entries for multi-player games, just under 50% mention communicating with other players through voice or text chat. It must be noted that participants were able to select both options in one entry: In fact, 17 of the multi-player gaming entries (18.1%) recorded simultaneous use of voice and text chat. It might seem surprising that voice chat was just as popular as written communication: One might expect learners to be shy to speak English or hesitate to connect with strangers over the internet. However, as one focus group participants explained, voice chat is so popular because game consoles and mobile devices (which are rivalling PCs as the most popular gaming platforms: Richter, 2018) make it difficult to use text chat, requiring players to painstakingly navigate to individual letters on a virtual keyboard, which is perceived as 'fiddly' and 'distracting' by the players. Video gaming with concurrent use of voice- or text-chat was found to be the activity with the highest levels of affective engagement, significantly above all other practices including gaming without player communication. These levels of enthusiasm might also serve to explain why participants are willing to use their L2 to engage with other players, and in fact spend significantly more time on this activity than on any other. In terms of cognitive engagement, video games and concurrent chatting were also rated highly and significantly above regular gaming. Interestingly though, both gaming categories ranked extremely low in terms of linguistic engagement levels (significantly above only listening to music). This begs the question whether video gaming is significantly linked to proficiency in this sample (as found in previous studies Sundqvist, 2009, 2011; Sundqvist & Sylvén, 2014), and if yes, what this means for the importance of conscious 'focus on form' in language acquisition.

F.3.7 Reading

As the whole category accounted only for 3.4% of the total recorded time with ISL practices, it may seem that, altogether, the participants in this study did not spend a lot of time on reading. The fact that very few students read books in their free time, much less in a second language, has been well documented through anecdotal evidence from parents and teachers, as well as in the ISLL literature (e.g. Olsson, 2011; Sundqvist, 2009; Sundqvist & Sylvén, 2014). However, it must be kept in mind that more than half of all social media use also involved reading (Section F.3.2.), as did other activities such as playing video games, watching videos with English subtitles, or communicating through text messages (see next section).

Coded in this category were diary entries pertaining to reading books, magazines or news articles, and online texts not on social media. In terms of frequency, online texts were most often read by the participants in this study (41.8% of all entries), followed by books (34.2%), magazines and news articles (13.9%).

Online texts were also read by almost twice as many of the participants (11.5%) than books (6.7%). However, the average time spent reading books ($M = 88.74, SD = 113.57$) was substantively longer than for online reading ($M = 48.61, SD = 63.80$), wherefore the latter category accounted for a much smaller proportion of the total duration of entries in this category (32.9% vs. 49.1%). 'Other' reading activities accounted for 10.1% of all entries and included reading comic books, instruction manuals, and English-language subtitles on a TV show voiced in another language (e.g. Japanese, Norwegian).

While reading, the participants reported average levels of affective engagement (not significantly different from most other activities), but significantly higher levels of cognitive engagement and linguistic engagement than in most other activities.

F.3.8 Listening

Finally, this category comprised activities which involved listening to spoken language, including overhearing — but not participating in — conversations in English and listening to podcasts, audiobooks, and radio programs. These activities accounted for the smallest number of all diary entries (30 entries, 2.6%) and amount of time spent on ISLL (1,160 minutes, 1.6%). Only 10.7% of the participants engaged in a listening activity at least once over the course of the study.

Just over half of the entries in this category concerned overhearing conversations that other people were having in English (1,160 minutes, 17 entries). Due to the nature of the diary survey, the circumstances in which the participants overheard these conversations are unknown, but in the focus groups some participants mentioned listening to family members speaking on the phone or on Skype to friends and family from abroad, as well as occasionally overhearing conversations of English-speaking tourists in public. It is unsurprising, then, that listening activities ranked lowest, together with playing music in the background, in terms of affective engagement, although the reported levels of cognitive and linguistic engagement were comparable to most other activities (and significantly above listening to music).

F.4 Gender and Types of Second Language Activities

Table F.6: Regression analyses, using gender to predict engagement in specific ISLPs

Activity	Gender → Activity (yes/no)							Gender → Activity (time spent)				
	<i>B</i>	<i>SE</i>	<i>p</i>	<i>Odds ratio</i>	R_N^2	n_{boys}	n_{girls}	<i>B</i>	<i>SE</i>	β	<i>p</i>	R^2
1. Films	- 1.91	.55	< .01	.15	.14	4	23	- 11.03	8.71	-.25	.217	.02
2. Serials or TV shows	-.91	.28	< .01	.40	.05	23	46	44.06	6.80	.00	.992	.00
3. Video-based social media	-.26	.23	.249	.75	.01	57	65	- 21.77	5.35	-.35	< .001	.11
								$F_{(1,120)} = 16.51, p < .001$				
4. Text-based social media	.05	.31	.865	1.05	.00	26	24	-.18	3.53	-.01	.960	.00
5. Mixed social media	-.80	.24	< .01	.45	.05	41	68	- 8.10	5.49	-.14	.143	.02
6. Writing	-.87	.36	< .05	.42	.04	13	27	- 2.57	5.61	-.07	.650	.01
7. Speaking	-.92	.33	< .01	.40	.04	16	34	- 12.69	7.26	-.25	.087	.04
8. Music (background)	-.46	.24	< .05	.63	.03	47	62	- 8.92	7.28	-.12	.223	.01
9. Music (focused)	- 1.21	.23	< .001	.30	.11	72	117	4.83	6.08	.06	.428	.00
10. Video games	1.01	.49	< .05	2.74	.05	16	6	- 31.11	26.33	-.26	.251	.05
11. Video games + text or voice chat	2.72	.74	< .001	15.28	.17	27	2	- 15.15	23.96	-.12	.532	.02
12. Reading	-.89	.32	< .01	.41	.04	17	35	11.23	5.66	.27	.053	.06
13. Listening	-.91	.28	< .01	.40	.05	23	46	.10	9.73	.00	.992	.00

F.5 Engagement Profiles and Types of Second Language Activities

Table F.7: Regression analyses, using engagement profile to predict engagement in specific ISLPs

Activity	Engagement Profile → Activity (yes/no)							Engagement Profile → Activity (time)				
	<i>B</i>	<i>SE</i>	<i>p</i>	<i>Odds</i>	R_N^2	n_{Grp1}	n_{Grp2}	<i>B</i>	<i>SE</i>	β	<i>p</i>	R^2
1. Films	.27	.41	.514	1.31	.00	15	12	-.76	8.79	-.02	.933	.00
2. Serials or TV shows	.46	.29	.105	1.59	.02	40	29	-32.14	9.03	-.40	<.01	.15
								$F_{(1,68)} = 12.68, p < .01$				
3. Video-based social media	.22	.25	.382	1.25	.00	64	58	-12.86	5.40	-.21	<.05	.04
								$F_{(1,121)} = 5.68, p < .05$				
4. Text-based social media	.18	.32	.586	1.19	.00	25	22	-4.46	3.65	-.18	.229	.01
5. Mixed social media	1.19	.27	<.001	3.28	.11	72	37	-21.58	5.06	-.38	<.001	.14
								$F_{(1,108)} = 18.21, p < .001$				
6. Writing	.14	.35	.690	1.15	.00	21	19	1.27	5.56	.04	.821	.00
7. Speaking	-.03	.31	.925	.97	.00	25	26	-13.58	7.09	-.26	.061	.05
8. Music (background)	.54	.26	.063	1.72	.02	64	48	-24.01	6.91	-.31	<.01	.09
								$F_{(1,111)} = 12.07, p < .01$				
9. Music (focused)	.34	.28	.234	1.40	.01	96	89	-23.15	5.93	-.28	<.001	.07
								$F_{(1,84)} = 15.27, p < .001$				
10. Video games	.71	.46	.122	2.03	.02	15	8	-17.42	27.05	-.14	.527	.01
11. Video games + text or voice chat	.57	.41	.158	1.77	.02	18	11	-58.60	21.71	-.46	<.05	.18
								$F_{(1,28)} = 7.28, p < .05$				
12. Reading	.41	.31	.193	1.51	.01	30	22	-4.54	5.92	-.11	.447	.01
13. Listening	.46	.29	.105	1.59	.02	40	29	-32.14	9.03	-.40	<.01	.15
								$F_{(1,68)} = 12.68, p < .01$				

Table F.8: Chi-squared Test comparing the popularity of activity types across the LPA groups

Activity	Number of participants who did engage in this activity		χ^2 (df = 1)	p	ω
	Group 1 (n = 126)	Group 2 (n = 128)			
1. Films			.43	.513	.04
2. Serials & TV shows			2.65	.137	.10
3. Video-based social media			.76	.382	.05
4. Text-based social media			.30	.586	.03
5. Mixed social media			20.67	< .001	.29
6. Writing			.16	.690	.03
7. Speaking			.01	.925	.01
8. Music (background)			4.55	.043	.13
9. Music (focused)			1.42	.233	.07
10. Video games			2.47	.116	.10
11. Video games & text or voice chat			2.03	.154	.09
12. Reading			1.71	.191	.08
13. Listening (spoken word)			2.65	.103	.10

Note: Greyed cells mark tests where a significant difference was found between the groups. A Bonferroni correction was applied to adjust for multiple comparisons, $\alpha = .0038$.

Table F.9: Mann-Whitney U tests comparing the time spent on different ISLPs across the LPA groups

Activity	Group 1		Group 2		U	z	p
	n	Mean rank	n	Mean rank			
1. Films	15	15.73	12	11.83	64.0	-1.30	.195
2. Serials & TV shows	40	41.69	29	25.78	312.5	-3.26	< .001
3. Video-based social media	64	69.97	58	52.16	1314.0	-2.78	.005
4. Text-based social media	25	24.25	22	23.73	269.0	- .13	.898
5. Mixed social media	72	64.39	37	36.73	656.0	-4.34	< .001
6. Writing	21	19.86	19	21.21	213.0	.37	.714
7. Speaking	25	29.24	26	22.88	244.0	-1.53	.126
8. Music (background)	64	63.12	48	47.68	1112.5	-2.49	.013
9. Music (focused)	96	104.65	89	80.43	3153.5	-3.08	.002
10. Video games	15	12.50	8	11.06	52.5	- .49	.626
11. Video games & text or voice chat	18	18.50	11	9.27	36.0	-2.84	.005
12. Reading	30	30.38	22	21.20	213.5	-2.17	.030
13. Listening (spoken word)	40	41.69	29	25.78	312.5	-3.26	.001

Note: A Bonferroni correction was applied to adjust for multiple comparisons, $\alpha = .0038$.

Appendix G. Findings Phase III, RQ 2: What are the Direct Relationships Between General Language Proficiency and Engagement in Informal Second Language Practices?

G.1 Proficiency and Types of Second Language Activities

Table G.1: Regression analyses, using prior proficiency to predict engagement in specific ISLPs

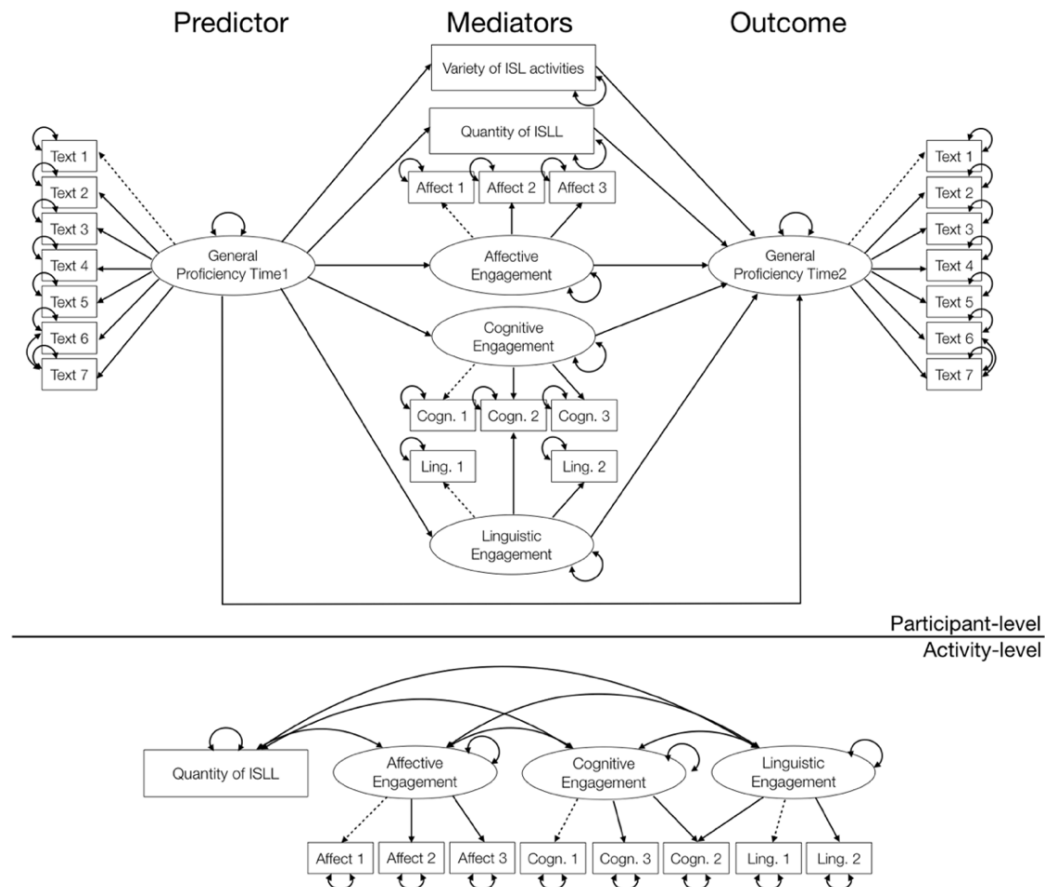
Activity	ProficiencyT1 → Activity (yes/no)					Proficiency T1 → Activity (time spent)					
	B	SE	p	Odds ratio	R _N ²	B	SE	β	p	R ²	n
1. Films	1.07	.36	< .01	2.92	.07	-.20	.70	-.06	.774	.00	27
2. Serials or TV shows	.74	.23	< .001	2.10	.05	.89	.51	.21	.085	.05	70
3. Video-based social media	.51	.19	< .01	1.67	.03	1.21	.43	.26	< .01	.07	124
						$F_{(1,111)} = 5.54, p < .05$					
4. Text-based social media	.30	.25	.236	1.35	.01	.96	.36	.36	< .05	.13	51
						$F_{(1,48)} = 5.81, p < .05$					
5. Mixed social media	.68	.20	< .01	1.97	.05	1.07	.42	.25	< .05	.06	110
						$F_{(1,101)} = 4.46, p < .05$					
6. Writing	.61	.28	< .05	1.84	.03	-.52	.59	-.15	.383	.02	40
7. Speaking	.23	.26	.380	1.26	.00	.12	.63	.03	.848	.00	51
8. Music (background)	.10	.19	.604	1.11	.00	.77	.46	.16	.098	.03	112
9. Music (focused)	.54	.19	< .01	1.71	.03	.34	.39	.07	.387	.00	190
10. Video games	-.22	.38	.573	.81	.00	-.09	1.52	-.01	.953	.00	23
11. Video games + text or voice chat	-.13	.33	.681	.88	.00	.55	1.21	.09	.649	.01	29
12. Reading	.29	.25	.250	1.34	.01	1.36	.50	.37	< .01	.14	53
						$F_{(1,47)} = 4.31, p < .01$					
13. Listening	.74	.23	< .01	2.10	.05	.89	.51	.21	.085	.05	70

Table G.2: Stepwise regression analyses, using engagement in specific ISLPS to predict changes in proficiency

Activity	ProficiencyT1 + Activity (yes/no) → ProficiencyT2				ProficiencyT1 + Activity (time spent) → ProficiencyT2				n
	B	SE	β	p	B	SE	β	p	
MODEL 1:	$R^2 = .87, F_{(1,330)} = 9211.14, p < .001$				$R^2 = .87, F_{(1,330)} = 9211.14, p < .001$				
ProficiencyT1	1.07	.01	.98	<.001	1.07	.01	.98	<.001	331
MODEL 2:	$R^2 = .87, F_{(1,330)} = 694.25, p < .001$				$R^2 = .87, F_{(1,330)} = 682.31, p < .001$				
ProficiencyT1	1.06	.01	.98	<.001	1.06	.01	.98	<.001	331
1. Films	-.03	.02	-.01	.181	-.01	.01	-.03	.476	27
2. Serials or TV shows	.02	.02	.01	.349	.01	.01	.03	.170	70
3. Video-based social media	.00	.02	.00	.959	-.01	.01	-.03	.089	124
4. Text-based social media	.00	.02	.00	.912	.02	.01	.05	.080	51
5. Mixed social media	.00	.02	.00	.898	.00	.01	.02	.930	110
6. Writing	.01	.02	.00	.686	-.01	.01	-.02	.627	40
7. Speaking	-.01	.02	-.01	.487	.00	.01	.01	.839	51
8. Music (background)	.00	.02	.00	.912	.00	.00	.00	.863	112
9. Music (focused)	.01	.02	.01	.668	.00	.00	.01	.741	190
10. Video games	.03	.03	.01	.381	.00	.01	.02	.638	23
11. Video games + text or voice chat	.01	.03	.01	.587	.00	.01	-.02	.625	29
12. Reading	.00	.02	.00	.921	.00	.01	-.01	.804	53
13. Listening	.02	.02	.01	.312	.01	.01	.03	.170	70

G.2 Proficiency and Dimensions of Engagement

Figure G.1: Engagement Mediation SEM, Full Latent Factor Structural Model

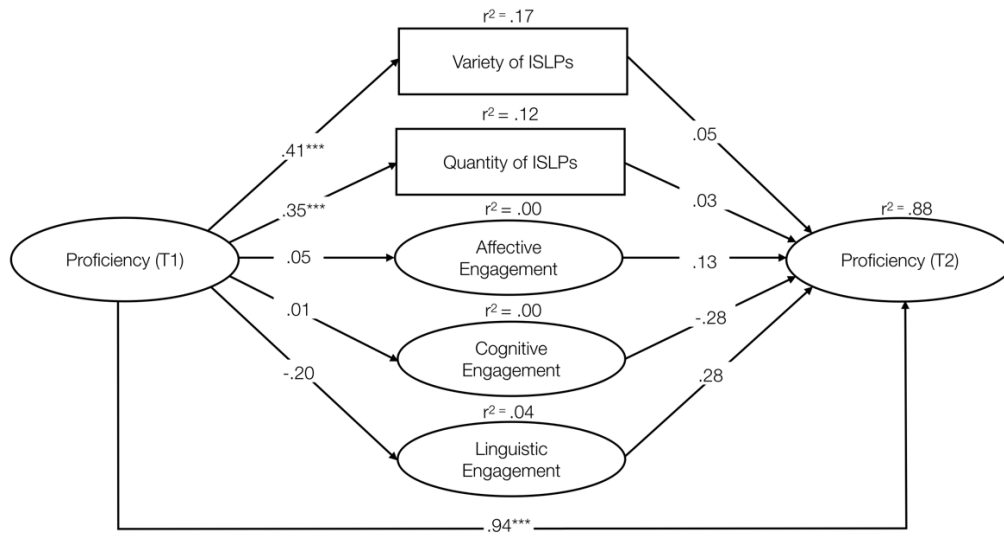


Note: The model also includes covariances between the latent and manifest engagement constructs at the participant-level. They were omitted from this illustration for visual clarity.

Table G.3: Proficiency and engagement Multi-Level Structural Equation Model. Significant regression path coefficients at the participant-level

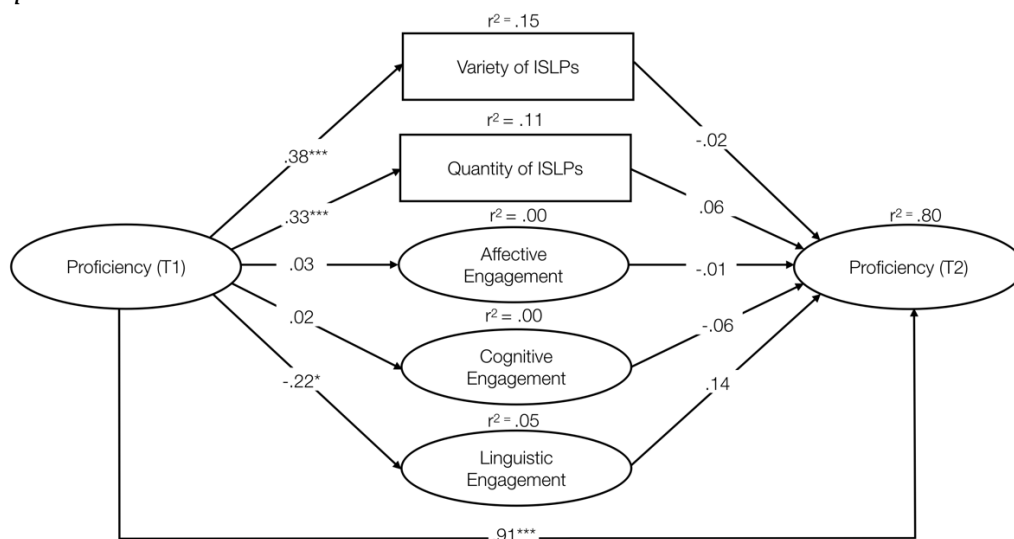
Predictors	Outcomes	B	SE	β	p
Proficiency T1	Variety of ISL activities	.58	.11	.41	< .001
Proficiency T1	Quantity of ISL	.54	.12	.35	< .001
Proficiency T1	Linguistic Engagement	-.19	.01	-.20	.052
Proficiency T1	Proficiency T2	1.13	.13	.94	< .001

Figure G.2: Proficiency and engagement Multi-Level Structural Equation Model (all participants). Regression paths at the participant-level



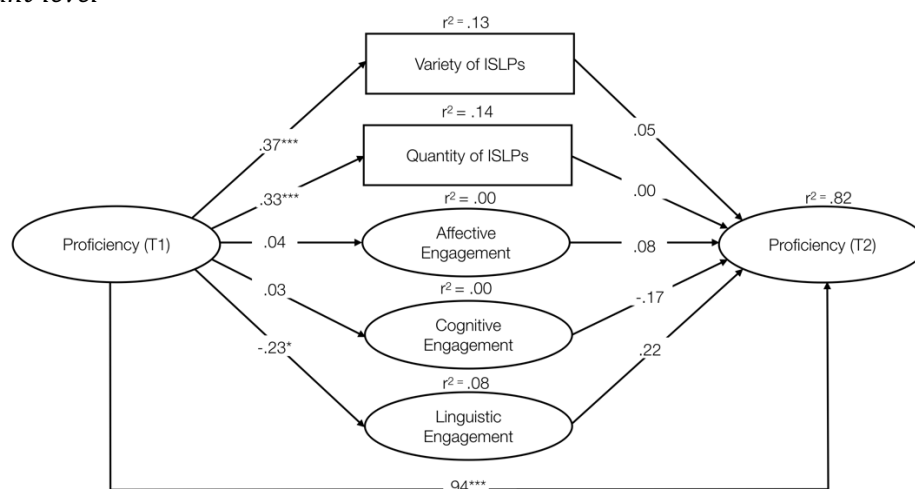
Estimation method: Robust Maximum Likelihood with listwise exclusion. $\chi^2_{[244]} = 443.34$, $p < .001$; $n_{obs} = 1113$, $n_{part} = 188$; $CFI = .96$, $TLI = .95$, $RMSEA = .03$, $SRMR_{w/i} = .04$, $SRMR_{b/w} = .07$.

Figure G.3: Proficiency and engagement Multi-Level Structural Equation Model (excluding participants who made diary entries over fewer than four days). Regression paths at the participant-level



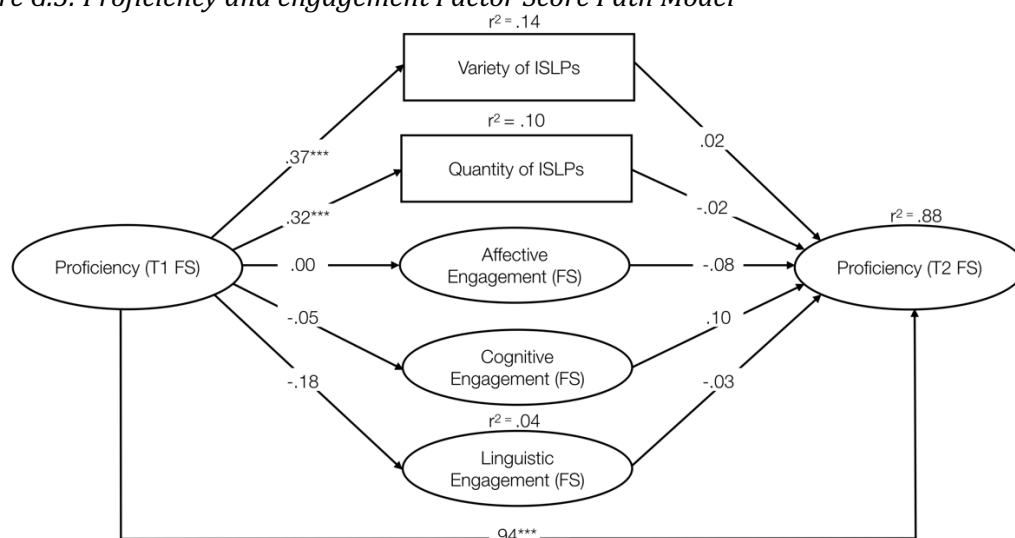
Estimation method: Robust Maximum Likelihood with listwise exclusion. $\chi^2_{[244]} = 423.95$, $p < .001$; $n_{obs} = 1043$, $n_{part} = 143$; $CFI = .96$, $TLI = .95$, $RMSEA = .03$, $SRMR_{w/i} = .04$, $SRMR_{b/w} = .07$.

Figure G.4: Proficiency and engagement Multi-Level Structural Equation Model (excluding participants who made diary entries over fewer than five days). Regression paths at the participant-level



Estimation method: Robust Maximum Likelihood with listwise exclusion. $\chi^2_{[244]} = 428.99$, $p < .001$; $n_{obs} = 963$; $n_{part} = 126$; $CFI = .96$, $TLI = .95$, $RMSEA = .03$, $SRMR_{w/i} = .04$, $SRMR_{b/w} = .07$.

Figure G.5: Proficiency and engagement Factor Score Path Model



Estimation method: Full Information Robust Maximum Likelihood. Fit indices were not available as this was a saturated model. $n = 349$.

Table G.4: Bivariate correlations between significant engagement variables and proficiency at T1 and T2

	Prior Proficiency (T1 FS)	Posterior Proficiency (T2 FS)
Overall engagement pattern (LPA group) ¹	-.38 **	-.37 **
Variety of ISL activities (FS)	.34 ***	.34 ***
Quantity of ISLL (FS)	.31 ***	.30 ***
Linguistic engagement (FS)	-.09	-.18

¹ As the LPA group was a binary variable, the values in this row represent point-biserial correlations.

Appendix H. Findings Phase III RQ 3: What are the Direct Relationships Between Language Learning Motivation and Attitudes Towards Language Learning and Engagement in Informal Second Language Practices?

H.1 Motivation, Attitudes and Dimensions of Engagement

Table H.1: Language learning orientations and engagement, Factor Score Path Model. Significant regression path coefficients.

<i>Predictors</i>	<i>Outcomes</i>	<i>B</i>	<i>SE</i>	<i>β</i>	<i>p</i>
PREDICTING ENGAGEMENT					
<i>Instrumental Obligation (T1)</i>	<i>Variety of ISLPs</i>	-.41	.18	-.41	< .05
<i>Instrumental Obligation (T1)</i>	<i>Quantity of ISLPs</i>	-.38	.19	-.38	< .05
<i>Knowledge Orientation (T1)</i>	<i>Quantity of ISLPs</i>	.22	.11	.22	< .05
<i>Intrinsic Motivation (T1)</i>	<i>Affective Engagement</i>	.28	.16	.28	.058
<i>Instrumental Obligation (T1)</i>	<i>Affective Engagement</i>	.68	.20	.68	< .01
<i>Knowledge Orientation (T1)</i>	<i>Affective Engagement</i>	.33	.12	.33	< .01
<i>Intrinsic Motivation (T1)</i>	<i>Cognitive Engagement</i>	.50	.14	.50	< .001
<i>Instrumental Obligation (T1)</i>	<i>Cognitive Engagement</i>	.72	.19	.72	< .001
<i>Knowledge Orientation (T1)</i>	<i>Cognitive Engagement</i>	.26	.12	.26	< .05
<i>Intrinsic Motivation (T1)</i>	<i>Linguistic Engagement</i>	.53	.15	.54	< .001
<i>Instrumental Obligation (T1)</i>	<i>Linguistic Engagement</i>	.49	.18	.48	< .01
PREDICTING CHANGE IN LANGUAGE LEARNING ORIENTATIONS (T2)					
<i>Intrinsic Motivation (T1)</i>	<i>Intrinsic Motivation (T2)</i>	1.30	.03	1.29	< .001
<i>Instrumental Obligation (T1)</i>	<i>Intrinsic Motivation (T2)</i>	.53	.03	.53	< .001
<i>International Orientation (T1)</i>	<i>Intrinsic Motivation (T2)</i>	.14	.01	.14	< .001
<i>Affective Engagement</i>	<i>Intrinsic Motivation (T2)</i>	-.21	.11	-.21	< .05
<i>Cognitive Engagement</i>	<i>Intrinsic Motivation (T2)</i>	.32	.15	.32	< .05
<i>Intrinsic Motivation (T1)</i>	<i>Instrumental Obligation (T2)</i>	-.13	.05	-.13	< .05
<i>Instrumental Obligation (T1)</i>	<i>Instrumental Obligation (T2)</i>	.86	.05	.85	< .001
<i>Knowledge Orientation (T1)</i>	<i>Instrumental Obligation (T2)</i>	.11	.01	.11	< .001
<i>Affective Engagement</i>	<i>Instrumental Obligation (T2)</i>	.33	.15	.33	< .05
<i>Cognitive Engagement</i>	<i>Instrumental Obligation (T2)</i>	-.49	.22	-.49	< .05
<i>Linguistic Engagement</i>	<i>Instrumental Obligation (T2)</i>	.26	.13	.26	< .05
<i>Instrumental Benefits (T1)</i>	<i>Instrumental Benefits (T2)</i>	.73	.03	.71	< .001
<i>International Orientation</i>	<i>Instrumental Benefits (T2)</i>	.11	.02	.11	< .001
<i>Intrinsic Motivation (T1)</i>	<i>Knowledge Orientation (T2)</i>	.09	.04	.09	< .05
<i>Instrumental Obligation (T1)</i>	<i>Knowledge Orientation (T2)</i>	.60	.04	.60	< .001
<i>Knowledge Orientation (T1)</i>	<i>Knowledge Orientation (T2)</i>	1.29	.03	1.27	< .001
<i>Intrinsic Motivation (T1)</i>	<i>International Orientation (T2)</i>	.12	.05	.12	< .05
<i>Instrumental Obligation (T1)</i>	<i>International Orientation (T2)</i>	.17	.05	.17	< .01
<i>Knowledge Orientation (T1)</i>	<i>International Orientation (T2)</i>	.08	.03	.08	< .01
<i>International Orientation (T1)</i>	<i>International Orientation (T2)</i>	.98	.02	.97	< .001

Table H.2: Second language motivational selves (current and future self) and engagement, Factor Score Path Model. Significant regression path coefficients

Predictors	Outcomes	B	SE	β	<i>p</i>
PREDICTING ENGAGEMENT					
Current Self (T1)	Variety of ISLPs	.23	.06	.23	< .001
Future Self (T1)	Variety of ISLPs	.17	.05	.17	< .01
Current Self (T1)	Quantity of ISLPs	.25	.06	.25	< .001
PREDICTING CHANGE IN CURRENT AND FUTURE MOTIVATIONAL SELVES (T2)					
Variety of ISLPs	Current Self (T2)	.08	.03	.08	< .05
Current Self (T1)	Current Self (T2)	.87	.03	.87	< .001
Variety of ISLPs	Future Self (T2)	.15	.05	.16	< .01
Future Self (T1)	Future Self (T2)	.62	.06	.63	< .001

Table H.3: Second language motivational selves (motivational distance) and engagement, Factor Score Path Model. Significant regression path coefficients

Predictors	Outcomes	B	SE	β	<i>p</i>
PREDICTING ENGAGEMENT					
Motivational Distance (T1)	Variety of ISLPs	-.08	.04	-.12	< .05
Motivational Distance (T1)	Quantity of ISLPs	-.13	.04	-.20	< .001
PREDICTING CHANGE IN MOTIVATIONAL SELF-DISTANCE (T2)					
Motivational Distance (T1)	Motivational Distance (T2)	.52	.03	.82	< .001
Affective Engagement	Motivational Distance (T2)	.42	.19	.43	< .05
Cognitive Engagement	Motivational Distance (T2)	-.49	.19	-.47	< .05
Linguistic Engagement	Motivational Distance (T2)	.36	.29	.36	< .05

Table H.4: Attitudes towards second language learning and engagement, Factor Score Path Model. Significant regression path coefficients

Predictors	Outcomes	B	SE	β	<i>p</i>
PREDICTING ENGAGEMENT					
Attitudes: Informal Learning (T1)	Variety of ISLPs	.30	.07	.30	< .001
Attitudes: Informal Learning (T1)	Quantity of ISLPs	.42	.07	.42	< .001
Attitudes: Informal Learning (T1)	Affective Engagement	.33	.08	.33	< .001
Attitudes: Formal Learning (T1)	Affective Engagement	.18	.08	.18	< .05
Attitudes: Informal Learning (T1)	Cognitive Engagement	.31	.09	.31	< .01
Attitudes: Formal Learning (T1)	Cognitive Engagement	.25	.08	.25	< .01
Attitudes: Formal Learning (T1)	Linguistic Engagement	.22	.09	.22	< .05
PREDICTING CHANGE IN ATTITUDES (T2)					
Attitudes: Informal Learning (T1)	Attitudes: Informal Learning (T2)	.88	.03	.89	< .001
Attitudes: Formal Learning (T1)	Attitudes: Formal Learning (T2)	.73	.04	.75	< .001

H.2 Motivation, Attitudes and Engagement Profiles

Table H.5: Motivation, attitudes, and engagement profiles, Factor Score Path Model. Significant regression path coefficients.

		B	SE	β	p
LANGUAGE LEARNING ORIENTATIONS					
No significant regression paths					
SECOND LANGUAGE MOTIVATIONAL SELVES: FUTURE AND CURRENT					
<i>Current Self (T1)</i>	<i>Engagement Profile</i>	-.17	.05	-.28	< .001
<i>Current Self (T1)</i>	<i>Current Self (T2)</i>	.77	.04	.83	< .001
<i>Future Self (T1)</i>	<i>Current Self (T2)</i>	.17	.04	.10	< .001
<i>Future Self (T1)</i>	<i>Future Self (T2)</i>	.72	.05	.68	< .001
SECOND LANGUAGE MOTIVATIONAL SELVES: MOTIVATIONAL DISTANCE					
<i>Motivational Distance (T1)</i>	<i>Engagement Profile</i>	.14	.05	.20	< .01
<i>Motivational Distance (T1)</i>	<i>Motivational Distance (T2)</i>	.71	.02	.82	< .001
ATTITUDES TOWARDS LANGUAGE LEARNING					
<i>Attitudes: Informal Learning (T1)</i>	<i>Engagement Profile</i>	-.28	.05	-.47	< .001
<i>Attitudes: Informal Learning (T1)</i>	<i>Attitudes: Informal Learning (T2)</i>	.82	.04	.83	< .001
<i>Attitudes: Formal Learning (T1)</i>	<i>Attitudes: Informal Learning (T2)</i>	-.15	.04	-.14	< .001
<i>Attitudes: Informal Learning (T1)</i>	<i>Attitudes: Formal Learning (T2)</i>	-.23	.05	-.23	< .001
<i>Attitudes: Formal Learning (T1)</i>	<i>Attitudes: Formal Learning (T2)</i>	.79	.04	.68	< .001
<i>Engagement Profile</i>	<i>Attitudes: Formal Learning (T2)</i>	-.23	.10	-.14	< .05
COMBINED					
<i>Current Self (T1)</i>	<i>Engagement Profile</i>	-.11	.05	-.18	< .05
<i>Attitudes: Informal Learning (T1)</i>	<i>Engagement Profile</i>	-.22	.06	-.36	< .01
<i>Attitudes: Informal Learning (T1)</i>	<i>Attitudes: Formal Learning (T2)</i>	-.22	.08	-.22	< .01
<i>Attitudes: Formal Learning (T1)</i>	<i>Attitudes: Formal Learning (T2)</i>	.79	.10	.68	< .001
<i>Engagement Profile</i>	<i>Attitudes: Formal Learning (T2)</i>	-.22	.10	-.13	< .05

Appendix I. Findings Phase III RQ 4: How do General Language Proficiency, Language Learning Motivation, and Attitudes Interact in Their Relationships with Engagement in Informal Second Language Practices?

I.1 Behavioural, Affective, Cognitive, and Linguistic Engagement

Table I.1: Proficiency, language learning orientations, and engagement, Factor Score Path Model. Significant regression path coefficients

<i>Predictors</i>	<i>Outcomes</i>	<i>B</i>	<i>SE</i>	<i>β</i>	<i>p</i>
PREDICTING LANGUAGE LEARNING ORIENTATIONS T1					
<i>Proficiency (T1)</i>	<i>Intrinsic Motivation (T1)</i>	.77	.08	.47	< .001
<i>Proficiency (T1)</i>	<i>Instrumental Obligation (T1)</i>	-.74	.08	-.45	< .001
<i>Proficiency (T1)</i>	<i>Knowledge Orientation (T1)</i>	.71	.08	.43	< .001
PREDICTING ENGAGEMENT					
<i>Proficiency (T1)</i>	<i>Variety of ISLPs</i>	.34	.11	.21	< .01
<i>Instrumental Obligation (T1)</i>	<i>Variety of ISLPs</i>	-.30	.07	-.30	< .001
<i>Proficiency (T1)</i>	<i>Quantity of ISLPs</i>	.31	.13	.19	< .05
<i>Knowledge Orientation (T1)</i>	<i>Quantity of ISLPs</i>	.27	.09	.27	< .01
<i>Intrinsic Motivation (T1)</i>	<i>Affective Engagement</i>	.46	.17	.46	< .01
<i>Instrumental Obligation (T1)</i>	<i>Affective Engagement</i>	.65	.20	.65	< .01
<i>Knowledge Orientation (T1)</i>	<i>Affective Engagement</i>	.44	.12	.44	< .001
<i>Proficiency (T1)</i>	<i>Cognitive Engagement</i>	-.37	.12	-.37	< .01
<i>Intrinsic Motivation (T1)</i>	<i>Cognitive Engagement</i>	.65	.14	.65	< .001
<i>Instrumental Obligation (T1)</i>	<i>Cognitive Engagement</i>	.64	.16	.64	< .001
<i>Knowledge Orientation (T1)</i>	<i>Cognitive Engagement</i>	.32	.08	.32	< .001
<i>Proficiency (T1)</i>	<i>Linguistic Engagement</i>	-.41	.12	-.25	< .01
<i>Intrinsic Motivation (T1)</i>	<i>Linguistic Engagement</i>	.61	.14	.61	< .001
<i>Instrumental Obligation (T1)</i>	<i>Linguistic Engagement</i>	.35	.14	.35	< .001
PREDICTING PROFICIENCY CHANGE (T2)					
<i>Proficiency (T1)</i>	<i>Proficiency (T2)</i>	1.60	.02	.98	< .001
PREDICTING LANGUAGE LEARNING ORIENTATIONS CHANGE (T2)					
<i>Proficiency (T1)</i>	<i>Intrinsic Motivation (T2)</i>	-.32	.16	-.20	< .05
<i>Intrinsic Motivation (T1)</i>	<i>Intrinsic Motivation (T2)</i>	1.40	.04	1.40	< .001
<i>Instrumental Obligation (T1)</i>	<i>Intrinsic Motivation (T2)</i>	.52	.07	.52	< .001
<i>Proficiency (T2)</i>	<i>Intrinsic Motivation (T2)</i>	.20	.10	.21	< .05
<i>Posterior Proficiency (T1)</i>	<i>Instrumental Obligation (T2)</i>	.41	.22	.25	< .05
<i>Intrinsic Motivation (T1)</i>	<i>Instrumental Obligation (T2)</i>	-.17	.06	-.17	< .05
<i>Instrumental Obligation (T1)</i>	<i>Instrumental Obligation (T2)</i>	.84	.09	.84	< .001
<i>Affective Engagement</i>	<i>Instrumental Obligation (T2)</i>	.26	.11	.26	< .05
<i>Cognitive Engagement</i>	<i>Instrumental Obligation (T2)</i>	-.39	.18	-.39	< .05
<i>Proficiency (T2)</i>	<i>Instrumental Obligation (T2)</i>	-.31	.14	-.31	< .05

Table I.2: Proficiency, motivational selves, attitudes towards language learning, and engagement, Factor Score Path Model. Significant regression path coefficients.

Predictors	Outcomes	B	SE	β	p
PREDICTING MOTIVATIONAL SELVES AND ATTITUDES (T1)					
Proficiency (T1)	Current Self (T1)	.87	.07	.53	< .001
Proficiency (T1)	Future Self (T1)	.48	.09	.30	< .001
Proficiency (T1)	Attitudes: Informal Learning (T1)	.53	.08	.32	< .001
PREDICTING ENGAGEMENT					
Proficiency (T1)	Variety in ISLPs	.37	.11	.23	< .01
Future Self (T1)	Variety in ISLPs	.13	.04	.13	< .01
Attitudes: Informal Learning (T1)	Variety in ISLPs	.14	.06	.14	< .01
Proficiency (T1)	Quantity in ISLL	.30	.13	.18	< .01
Attitudes: Informal Learning (T1)	Quantity in ISLL	.30	.06	.30	< .001
Attitudes: Informal Learning (T1)	Affective Engagement	.31	.09	.32	< .01
Attitudes: Formal Learning (T1)	Affective Engagement	.21	.09	.21	< .05
Proficiency (T1)	Cognitive Engagement	-.19	.10	-.12	< .05
Attitudes: Informal Learning (T1)	Cognitive Engagement	.21	.06	.22	< .01
Attitudes: Formal Learning (T1)	Cognitive Engagement	.21	.08	.22	< .01
Proficiency (T1)	Linguistic Engagement	-.19	.10	-.12	< .05
Attitudes: Formal Learning (T1)	Linguistic Engagement	.14	.07	.14	< .05
PREDICTING PROFICIENCY CHANGE (T2)					
Proficiency (T1)	Proficiency (T2)	1.60	.02	.94	< .001
PREDICTING CHANGE IN MOTIVATIONAL SELVES AND ATTITUDES (T2)					
Proficiency (T1)	Current Self (T2)	-.66	.25	-.40	< .01
Current Self (T1)	Current Self (T2)	.81	.04	.81	< .001
Proficiency (T2)	Current Self (T2)	.47	.15	.47	< .01
Proficiency (T1)	Future Self (T2)	-.71	.37	-.71	< .05
Variety of ISLPs	Future Self (T2)	.07	.04	.07	< .05
Future Self (T1)	Future Self (T2)	.65	.08	.65	< .001
Proficiency (T2)	Future Self (T2)	.53	.04	.53	< .05

Table I.3: Proficiency, motivational self-distance, attitudes towards language learning, and engagement, Factor Score Path Model. Significant regression path coefficients

Predictors	Outcomes	B	SE	β	p
PREDICTING MOTIVATIONAL SELF-DISTANCE AND ATTITUDES (T1)					
Proficiency (T1)	Motivational Distance (T1)	-.61	.07	-.37	< .001
Proficiency (T1)	Attitudes: Informal Learning (T1)	.53	.08	.32	< .001
PREDICTING ENGAGEMENT					
Proficiency (T1)	Variety in ISLPs	.53	.09	.32	< .001
Attitudes: Informal Learning (T1)	Variety in ISLPs	.14	.06	.14	< .05
Proficiency (T1)	Quantity in ISLPs	.35	.12	.21	< .01
Attitudes: Informal Learning (T1)	Quantity in ISLPs	.30	.06	.30	< .001
Attitudes: Informal Learning (T1)	Affective Engagement	.31	.09	.32	< .01
Attitudes: Formal Learning (T1)	Affective Engagement	.21	.09	.21	< .05
Proficiency (T1)	Cognitive Engagement	-.19	.09	-.12	< .05
Attitudes: Informal Learning (T1)	Cognitive Engagement	.21	.06	.22	< .01
Attitudes: Formal Learning (T1)	Cognitive Engagement	.21	.08	.22	< .01
Proficiency (T1)	Linguistic Engagement	-.19	.10	-.12	< .05
Attitudes: Formal Learning (T1)	Linguistic Engagement	.14	.07	.14	< .05
PREDICTING PROFICIENCY CHANGE (T2)					
Proficiency (T1)	Proficiency (T2)	1.60	.02	.94	< .001
PREDICTING CHANGE IN MOTIVATIONAL SELVES AND ATTITUDES (T2)					
Motivational Distance (T1)	Motivational Distance (T2)	.79	.05	.79	< .001
Attitudes: Informal Learning (T1)	Motivational Distance (T2)	-.13	.05	-.13	< .01
Affective Engagement (T1)	Motivational Distance (T2)	.41	.19	.40	< .05
Cognitive Engagement (T1)	Motivational Distance (T2)	-.57	.29	-.56	< .05
Linguistic Engagement (T1)	Motivational Distance (T2)	.33	.17	.33	.058

I.2 Engagement Profiles

Table I.4: Proficiency, motivation, attitudes, and engagement profiles, Factor Score Path Model. Significant regression path coefficients.

<i>Predictors</i>	<i>Outcomes</i>	<i>B</i>	<i>SE</i>	<i>β</i>	<i>p</i>
<i>Proficiency (T1)</i>	<i>Current Self (T1)</i>	1.44	.18	.53	< .001
<i>Proficiency (T1)</i>	<i>Attitudes: Informal Learning (T1)</i>	.82	.21	.30	< .001
<i>Proficiency (T1)</i>	<i>Engagement Profile</i>	-.66	.18	-.34	< .001
<i>Attitudes: Informal Learning (T1)</i>	<i>Engagement Profile</i>	-.22	.07	-.32	< .01
<i>Attitudes: Informal Learning (T1)</i>	<i>Attitudes: Formal Learning (T2)</i>	-.23	.05	-.22	< .001
<i>Attitudes: Formal Learning (T1)</i>	<i>Attitudes: Formal Learning (T2)</i>	.77	.06	.67	< .001
<i>Engagement Profile</i>	<i>Attitudes: Formal Learning (T2)</i>	-.17	.09	-.11	< .05