

Individual differences in early instructed language learning

The role of language aptitude,
cognition, and motivation

Edited by

Raphael Berthele

Isabelle Udry

Eurosla Studies 5



EuroSLA Studies

Editor: Amanda Edmonds, Université Côte d'Azur

Associate editors:

Gabriele Pallotti, University of Modena and Reggio Emilia

Ineke Vedder, University of Amsterdam

In this series:

1. Pérez Vidal, Carmen, Sonia López-Serrano, Jennifer Ament & Dakota J. Thomas-Wilhelm (eds.). Learning context effects: Study abroad, formal instruction and international immersion classrooms
2. Saturno, Jacopo. Utterance structure in initial L2 acquisition.
3. Bardel, Camilla & Laura Sánchez (eds.). Third language acquisition: Age, proficiency and multilingualism.
4. Edmonds, Amanda, Pascale Leclercq & Aarnes Gudmestad (eds.), Interpreting language-learning data.
5. Berthele, Raphael & Isabelle Udry (eds.). Individual differences in early instructed language learning: The role of language aptitude, cognition, and motivation.

Individual differences in early instructed language learning

The role of language aptitude,
cognition, and motivation

Edited by

Raphael Berthele

Isabelle Udry

Raphael Berthele & Isabelle Udry (eds.). 2021. *Individual differences in early instructed language learning: The role of language aptitude, cognition, and motivation* (Eurosla Studies 5). Berlin: Language Science Press.

This title can be downloaded at:

<http://langsci-press.org/catalog/book/313>

© 2021, the authors

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):

<http://creativecommons.org/licenses/by/4.0/> 

ISBN: 978-3-96110-324-9 (Digital)

978-3-98554-020-4 (Hardcover)

ISSN: 2626-2665

DOI: 10.5281/zenodo.5378471

Source code available from www.github.com/langsci/313

Collaborative reading: paperhive.org/documents/remote?type=langsci&id=313

Cover and concept of design: Ulrike Harbort

Proofreading: Adam Stone, Alexandr Rosen, Alys Boote Cooper, Amir Ghorbanpour, Andreas Hölzl, Annika Schiefner, Brett Reynolds, Claudia Marzi, Craevschi Alexandru, Elen Le Foll, Eliane Lorenz, Esther Yap, Ikmi Nur Oktavianti, Jean Nitzke, Jeroen van de Weijer, Leonie Twente, Marten Stelling, M. Chiara Miduri

Fonts: Libertinus, Arimo, DejaVu Sans Mono

Typesetting software: Xe_{La}T_EX

Language Science Press

xHain

Grünberger Str. 16

10243 Berlin, Germany

langsci-press.org

Storage and cataloguing done by FU Berlin

Freie Universität



Berlin

Contents

	Introduction to the volume	
	Isabelle Udry & Raphael Berthele	iii
1	Language Aptitude at Primary School (LAPS): Theoretical framework of the project	
	Isabelle Udry, Raphael Berthele & Carina Steiner	1
2	Language Aptitude at Primary School (LAPS): Research design	
	Carina Steiner, Raphael Berthele & Isabelle Udry	51
3	The smart, the motivated and the self-confident: The role of language aptitude, cognition, and affective variables in early instructed foreign language learning	
	Isabelle Udry & Raphael Berthele	71
4	Predicting L2 achievement: Results from a test battery measuring language aptitude, general learning ability, and affective factors	
	Raphael Berthele, Jan Vanhove, Carina Steiner, Isabelle Udry & Hansjakob Schneider	91
5	Dispositions for language learning and social differences	
	Raphael Berthele	105
6	Creative thinking as an individual difference in task-based language teaching and learning	
	Isabelle Udry	125
7	The closer the better? Investigating L2 motivation of young learners in different contexts	
	Carina Steiner	143
8	The dynamics of young learners' L2 motivation: A longitudinal perspective	
	Carina Steiner	163

9	Language aptitude in German as a school language and English as a foreign language in primary school	
	Hansjakob Schneider	179
10	The stability of language aptitude: Insights from a longitudinal study on young learners' language analytic abilities	
	Isabelle Udry & Jan Vanhove	197
11	Summing up: Individual differences in primary school foreign language learning	
	Raphael Berthele & Isabelle Udry	211
	Index	225

Introduction to the volume

Isabelle Udry^{a,b} & Raphael Berthele^a

^aUniversity of Fribourg, Institut de Plurilinguisme ^bZurich University of Teacher Education

This introduction outlines the main focus and features of the project Language Aptitude at Primary School (LAPS). We begin with the rationale for the study and some clarification on terminology used throughout the book. Next, we discuss key concepts underlying language learning ability and early foreign language tuition. Finally, we provide an overview of the study design and the contents of the volume.

Since the beginning of the new millennium, early foreign language teaching and learning has seen important changes, namely the lowering of the starting age for language classes across Europe and the mandatory introduction of two foreign languages at primary schools in Switzerland where this study took place. These developments incited controversy and led to the need for empirical evidence that could underpin the arguments. It was against this backdrop that the project *Language Aptitude at Primary School* (LAPS) emerged. Our intention was to provide new insights into what shapes 10- to 12-year-old children's foreign language learning in minimal input settings with 2–3 weekly lessons. To this aim, we assessed the impact of a set of individual difference (ID) variables and environmental factors on young learners' developing foreign language proficiency over a period of two academic years. Particular attention was paid to language aptitude, a construct that has been extensively researched with adults, but has only recently sparked scholarly interest in relation to young learners (for a discussion see Chapter 1, §2.2). The results gathered from a range of cross-sectional and longitudinal analyses will be presented in this volume.

1 Reader's guide

This introduction contains all the information needed to follow the empirical chapters. In addition, two introductory chapters provide more detail on the theoretical framework of the LAPS project (Chapter 1) and the study design (Chapter 2). Readers are invited to read Chapters 1 and 2 before embarking on the rest



Isabelle Udry & Raphael Berthele. 2021. Introduction to the volume. In Raphael Berthele & Isabelle Udry (eds.), *Individual differences in early instructed language learning: The role of language aptitude, cognition, and motivation*, iii–xvii. Berlin: Language Science Press. DOI: 10.5281/zenodo.5464739



of the volume or consult them as questions arise during reading. Chapters 3 to 10 cover different aspects of the LAPS project (outlined in §6 of this chapter) and are conceived as independent texts, with the main information being summarized in the abstracts and methodology sections of each chapter. For the sake of replicability, supplementary material, including datasets and R scripts, have been made available online: <https://osf.io/hstv7/>.

With four official languages (German, French, Italian, and Romansh) and a variety of heritage languages, Switzerland's linguistic landscape is certainly diverse. This calls for some introductory remarks on the use of terminology in this volume.

L1 refers to the first language of the children. School language German (or German as a school language) describes the language of literacy or language of instruction in project region. Second language (L2) and third language (L3) designate the foreign languages taught at primary school in order of introduction: L2 refers to the first foreign language and L3 to the second foreign language introduced as part of the mandatory Swiss curriculum.

We are aware that on entering primary school, many children in Switzerland already have several languages in their repertoire, either because they are heritage language speakers, because they speak a Swiss German (Alemannic) dialect at home, or because of family ties with other linguistic regions of the country (see Berthele 2020 for a discussion of these difficulties in counting languages in the multilingual repertoire). To these children, foreign languages taught at school are actually their fourth or fifth language and German may not be their L1. Nevertheless, we adhere to using L2/L3 for instructed language teaching and learning, particularly for ease of reading.

As will be outlined in §5.2, the project consists of two subprojects, LAPS I and LAPS II. Throughout the volume, we will use the term LAPS when referring to the project in general, and LAPS I or LAPS II when talking about the specific subprojects.

2 A talent for language learning

Being a successful language learner often comes with a great deal of recognition. Whether it be the hyper-polyglot conversing fluently in many languages (e.g. Erard 2012), or the person who has picked up a native-like accent in a language different from their first (Flege & Mackay 2011, Christiner & Reiterer 2015), both are likely to encounter admiration for their achievements, and most certainly the question: "How do you do it?"

The notion of a talent for language learning was first theorized in the United States by John B. Carroll during the 1950s and 60s. The main reason for studying the characteristics of successful language learners was to provide government institutions with tools to select promising candidates for state-funded language courses. To this aim, Carroll (1964) administered a range of tests deemed to capture key abilities for language learning to members of staff at the US Army. From the results, he derived four language-related factors he subsumed under the term *language aptitude*:

1. *phonetic coding* (the ability to store, identify, and remember auditory phonetic material),
2. *grammatical sensitivity* (the ability to recognize the grammatical functions of elements in clauses),
3. *inductive language learning* (the ability to discover grammatical rules independently), and
4. *rote memory* (the ability to memorize new words rapidly and then retrieve them from memory).

Based on these components, Carroll & Sapon (1959) developed the Modern Language Aptitude Test (MLAT) which became widely used for selection and research purposes. However, the view on language aptitude as a predetermined attribute that could regulate access to language education soon came under scrutiny by educational stakeholders and scholars. Also, new (communicative) approaches to language teaching were considered to transform learning in a way that neutralized individual differences in language learning aptitude (Skehan 2002: 72). Concomitant with dominant views on individuals and societies in academia in the last decades of the 20th century, the idea that people differ in their ability to think and learn beyond what can be explained by social differences had become very unfashionable, to say the least. As argued in Pinker (2003: 28) the idea of the “ghost in the machine”, that is that humans are malleable and can be made better (or worse) by pedagogy became the “watchword of social science”.

2.1 New perspectives

While the discomfort with the Carrollian aptitude construct led to a marked decrease in scientific activity for several decades, language aptitude never entirely

disappeared from the research agenda. Recent scholarly interest has moved away from merely forecasting L2 achievement for selective purposes. Instead, relating language aptitude to SLA theories has become a main focus that has drawn attention from disciplines beyond applied linguistics, such as educational and cognitive psychology, or the neurosciences (Wen, Skehan, Sparks, et al. 2019).

Extending on the cognitive-linguistic focus reflected in the early stages of aptitude research, the ability to learn and communicate in a foreign language is currently regarded as being governed by a multitude of factors which can be grouped into three categories (Reiterer 2009): *biological* (e.g., DNA, sex, hormones), *linguistic/socio-cultural* (e.g., quality and quantity of input, language attitudes, typological distance/closeness between languages), and *psycho(bio)logical factors* (e.g., motivation, verbal intelligence, and language aptitude as defined in the previous paragraphs). A broad view that subsumes biological, language-related, cognitive, and affective factors that are studied from multiple scientific perspectives, holds promising prospects for advancing theories of foreign language learning and SLA. Recently, a number of innovative research projects have been conducted, the results of which can be consulted for instance in volumes by Reiterer (2019) or Wen, Skehan, Biedroń, et al. (2019).

2.2 Nature and nurture

Reviewing various studies that defined language aptitude as the ability to deal with language phonetically, grammatically, lexically or pragmatically, Reiterer (2019) concludes that these skills and abilities are normally distributed in the population. With reference to the bell-shaped curve, this means that a small group of about 15% will achieve very high, possibly near-native proficiency, while another 15% will retain very little of a foreign language. The remaining majority of about 70% will reach average skill levels. Language learning ability is therefore present in all individuals to varying degrees and the question of language talent cannot be answered by a simple yes or no statement.

Normally distributed characteristics, such as height, weight or intelligence, have been linked to some biological underpinnings (Reiterer 2019). There has been ongoing debate in psycholinguistics on the extent to which language learning and variation in achievement are genetically wired. Recent large-scale adoption and twin studies provide evidence that a considerable proportion of success in second and foreign language learning can be explained by hereditary factors. According to some studies, the genetic-makeup explains 50% or more of the variance in various aspects of human cognition (Dale et al. 2010, Stromswold 2001,

Rimfeld et al. 2015). This would still leave up to half of the variance to be attributed to factors other than genes, an observation that may alleviate some of the early apprehensions about language aptitude being fixed at birth and paving the way for egalitarian practices in education. It also ties in with the question of whether language learning ability could be influenced or even be trained by providing specific educational conditions.

In sum, key questions regarding language learning ability are a) the impact of individual predispositions (including aptitude, general learning abilities, and motivation) and external influences (such as socioeconomic status, teaching conditions, quantity and quality of input) on language competence; b) the extent to which these influencing factors can be changed by experience or training; and c) the relationship between individual predispositions, especially domain specific and general cognitive abilities.

3 Children and foreign language learning in Europe ...

The European Union (EU) considers linguistic and cultural diversity as one of its main assets worth promoting. Based on recommendations made by the Barcelona European Council (2002: 19), the general aim for EU citizens is now mastery of basic skills in at least two foreign languages. An early start to language learning at school has been declared a key strategy in pursuing this ambitious objective (European Commission 2004). This has led to the starting age for foreign language classes being lowered across Europe in recent years. According to the 2012 Euridyce/Eurostat survey conducted in 32 European countries, the usual starting age in 2009/10 was between 6 and 9 years. 78% of all children attending primary school in 2009/10 were learning a foreign language, in most cases English (Euridyce & Eurostat 2012: 10f).

The introduction of early language teaching in Europe and beyond has not gone without some major challenges, particularly in relation to developing appropriate educational frameworks. Major difficulties emerged in drafting generalizable policies underpinned by sound assumptions about children's learning (Johnstone 2009) and implementing these policies with adequate resources, such as age-appropriate teaching models and materials, or well-prepared teachers (for a discussion see Garton et al. 2011). Early instructed language learning also led to increased research activity, with teaching principles and age-related questions being explored in several large-scale studies, most notably by Edelenbos et al. (2006), Muñoz (2006), Nikolov & Csapó (2010), Enever (2011), Garton et al. (2011), Pfenninger (2016), Jaekel et al. (2017), and Baumert et al. (2020).

4 ... and Switzerland

The European trend has no doubt influenced policy development in Switzerland. Because of its multilingual context with four official languages, foreign language learning has a longstanding tradition in the country. As early as 1975, the Swiss government's recommendation for teaching one foreign language at primary school was being implemented throughout the country (for more details on the history of foreign language teaching in Switzerland see Giudici & Grizelj 2016). In the early 2000s, a new national strategy prescribed the introduction of even two foreign languages at primary school (EDK 2004), one at age 9, the second at age 11. At least one of them had to be a national language, the other could be English. Owing to the federal system, the cantons were free to choose how they would put the strategy into action, i.e. which two languages they wanted to introduce to children in what order. This led to considerable debate, as some cantons opted to start with English, rather than a national language. This choice was seen as a threat to national cohesion by some citizens, especially speakers of the minority national languages French, Italian and Romansh (Stotz 2006).

Moreover, concerns were expressed about some learner groups being overwhelmed by the demands of studying two foreign languages. However, while heavily debated, it was difficult to substantiate these fears with empirical evidence.¹ In the end, and as for many aspects of educational planning, the cantons were left to handle dispensation from foreign language classes as they saw fit.

5 The project Language Aptitude at Primary School (LAPS)

The project comprises two parts, LAPS I and LAPS II, which took place between spring 2017 and spring 2019. Samples and data collection are summarized in Tables 1 and 2. The children came from Swiss public schools, i.e. non-selective state-funded schools that teach all children living in their catchment area. Participants in both projects attended grades 4 and 5 (10 and 11 years) at the beginning of the study and were learning an L2 and L3 with 2–3 weekly lessons per language as part of the mandatory curriculum. At the beginning of the study, all participants completed a test battery assessing a great number of individual difference (ID) variables (see Figure 1). The results were related to their L2 and/or L3 proficiency. In the first part of the project (LAPS I), we considered L2 French and L3 English proficiency cross-sectionally ($n = 174$). In the second subproject (LAPS

¹The Schweizerische Akademie der Geisteswissenschaften (2015) published an overview on the arguments used in the Swiss debate.

II, $n = 637^2$), we recorded children's development of L2 English proficiency and school language German over two academic years (1.5 years).

5.1 Individual difference (ID) variables and environmental factors

Starting from the assumption that learning in general is influenced by a multitude of individual and contextual factors, we adopted a largely psycholinguistic perspective for this study with reference to the literature on individual differences (ID) in foreign language learning. We also included variables pertaining to the children's social background as previous research has consistently found them to be related to learning. A main objective of the project was to better understand how language aptitude in the Carrollian sense is implicated in child learning, an issue that has received little attention in research so far. Independent variables selected for the study fall into four categories:

1. Language aptitude
 - grammatical sensitivity
 - inductive learning³
 - phonetic coding ability
 - rote memory
2. General cognitive abilities or general learning abilities
 - intelligence
 - working memory
 - creativity
 - cognitive style (field independence)
3. Affective Dispositions
 - L2/L3 motivation
 - foreign language learning anxiety
 - L2/L3 self-concepts
 - dedication

²This number pertains to the total of individuals participating in at least one of three data collections of LAPS II. Due to children leaving or joining the project, this number differs from the total for each data collection indicated in Table 2.

³Based on a definition by Skehan (1998), grammatical sensitivity and inductive ability can be subsumed as language analytic ability.

- perceived support from teachers and parents
- locus of control

4. Environmental factors or background variables

- socioeconomic status (SES)
- language background
- teaching paradigm

Figure 1 shows the structure of the independent variables. Environmental factors are assumed to be overarching, as it is difficult for the individual to change them. Individual predispositions (or ID variables) are nested within these environmental factors. Based on the literature, it is assumed that there is interaction between social status, linguistic background or approaches to teaching, and the affective dispositions, such as motivation to learn foreign languages and anxiety. The dynamicity between these categories is indicated by the dotted line (and the arrow pointing from environmental to affective). Also, some fluidity between language aptitude and general cognitive variables is expected, most notably for memory functions (see Chapter 1, §2.3 for a discussion). Rote memory which stands for the ability to rapidly map meaning to sound/word form, is part of the aptitude construct. Recently however, some researchers have suggested extending this component with a more current definition of memory, based on the working memory model by Baddeley & Hitch (1974).

5.2 Research questions

The following research questions were addressed in the LAPS project:

- What ID variables are predictive of children's L2 proficiency and to what extent?
- What is the relationship between these variables, especially aptitude and general learning abilities?
- What developmental patterns can be observed in L1 and L2 proficiency, aptitude, and motivation?
- How do environmental factors affect children's L2 learning, (most notably SES, living close to a native speaking community)?

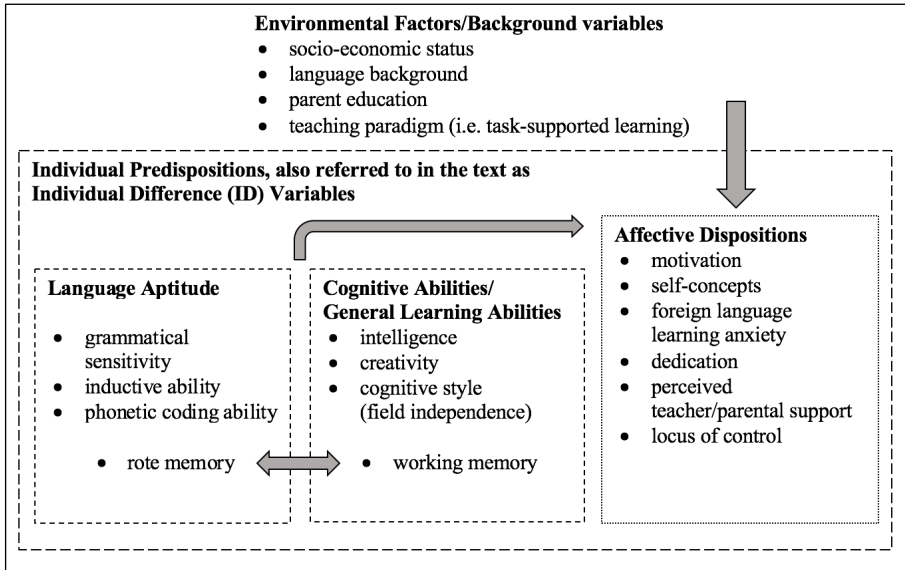


Figure 1: Structure of independent variables. Dotted lines indicate that the clear-cut categorization can be questioned. Arrows show expected direction of interaction.

5.3 Design and procedures

5.3.1 LAPS I

The first subproject was conducted with 4th and 5th graders from 10 classes located at the border with French-speaking Switzerland. Children's school language was German, they learnt L2 French (starting in 3rd grade, at 9 years old) and L3 English (starting in 5th grade, at 11 years old). Two data collections took place: T1 in spring 2017 ($n = 174$, mean age 11.1) and T2 in spring 2018 ($n = 158$, mean age 12.1).

In LAPS I, the test battery was piloted. Subsequently, minor changes were made for LAPS II (see Chapter 2, §3 for details). A second data collection T2 was included for two reasons: 1) to investigate the longitudinal development of affective dispositions. We wanted to find out how living close to native speakers of French would be reflected in the children's motivation to learn French and English over time (Chapter 7); 2) to understand the relationships between L2 and L3 skills (published in Berthele & Udry 2019). To address these issues, the questionnaire on affective dispositions was re-administered and a measure of L3 English proficiency was added at T2. T2 had not been part of the overall design and was

added as a follow-up project in the context of research training for students in the Fribourg multilingualism Master's program.

Table 1: Summary for main information LAPS I

Date	Participants	Independent variables	Language proficiency
T1 spring 2017	4 th /5 th grade <i>n</i> = 174 mean age: 11.1	Entire test battery: aptitude general cognitive abilities affective ID variables background variables	L2 French School language German
T2 spring 2018	5 th /6 th grade <i>n</i> = 158 mean age: 12.1	L2/L3 affective ID variables	L3 English

5.3.2 LAPS II

32 classes from the Eastern part of Switzerland participated in LAPS II for a period of two academic years (1.5 years in total). At the beginning of the study, the children were either in 4th or 5th grade (mean age 10.5), at the end of the study in 5th or 6th grade (12.1 years old). These children's school language was German, they learnt L2 English (starting in 2nd grade, at the age of 8) and L3 French (starting in 5th grade, at the age of 11).

LAPS II was longitudinal so we could trace the development of a) language proficiency in L2 English, b) school language German, c) language aptitude (grammatical sensitivity and inductive ability), d) affective dispositions.

Data were collected three times in the same classes: At T1 (autumn 2017), we administered the entire test battery with all ID variables, L2 proficiency, and proficiency in school language German to all children. At T2 (spring 2018) and T3 (spring 2019), five measures were re-administered to the same participants to monitor longitudinal development: 1) L2 English proficiency, 2) school language German proficiency, 3) language aptitude (grammatical sensitivity), 4) language aptitude (inductive ability), 5) L2/L3 motivation questionnaire.

6 Findings

The results of the project are presented in Chapters 3 to 10. Chapter 3 discusses the various dimensions of the ID variables assessed in the test battery and their

Table 2: Summary main information for LAPS II

Date	Participants	Independent variables	Language proficiency
T1 autumn 2017	4 th /5 th grade <i>n</i> = 615 mean age: 10.5	entire test battery: aptitude general cognitive abilities affective ID variables background variables	L2 English school language German proficiency
T2 spring 2018	4 th /5 th grade <i>n</i> = 578 mean age: 11.1	aptitude: grammatical sensitivity inductive ability L2/L3 affective ID variables	L2 English school language German proficiency
T3 spring 2019	5 th /6 th grade <i>n</i> = 566 mean age: 12.1	aptitude: grammatical sensitivity inductive ability L2/L3 affective ID variables	L2 English school language German proficiency

influence on L2 learning by primary school children. Chapter 4 deals with the predictive power of these ID variables for the participants' L2 proficiency.

The second part of the volume is devoted to more specific issues of the LAPS project. Chapter 5 examines the impact of socioeconomic factors, Chapter 6 looks into a less researched variable, creativity, within the context of task-based language learning, and Chapter 7 is dedicated to the role of motivation for L2/L3 learning at primary school. Chapters 8 to 10 address developmental patterns associated with ID variables over two academic years. Chapter 8 investigates changes in motivation, Chapter 9 covers the relationship between skills in the school language German and L2 English proficiency, and Chapter 10 explores the extent to which language aptitude, i.e. its language analytic subcomponent, remains stable over time.

We hope that this volume will incite discussion on early instructed language learning and encourage further scientific activity related to child L2/L3 learning, which we deem to be a viable research topic.

Acknowledgments

The LAPS project has been funded by the Research Centre on Multilingualism at the University of Fribourg and Teacher Training College of Fribourg, Switzerland.

Many people have contributed to the successful execution of the project. Most importantly, we thank the teachers and pupils for their commitment. This project would have been impossible without them. Many thanks to Charles W. Stansfield for letting us translate and adapt forms of the MLAT-E and PLAB tests. Our thanks go to a panel of experts who have guided us with their invaluable advice throughout the entire endeavour: Esther Geva, Joachim Grabowski, Susanne Reiterer. We are grateful to Amelia Lambelet for her contribution to LAPS I. To Peter Lenz for generously sharing his expertise and assisting us in selecting a suitable English measure. We thank Raphael Marguet from the Atelier Multimédia at the PH Fribourg for his support in recording test instructions. We would also like to acknowledge the time and effort devoted by three anonymous reviewers to improving the quality of our manuscript. Last but not least, we are thankful to a group of dedicated fieldworkers for their help with data collection and processing: Josef Adler, Thomas Aeppli, Nael Ackermann, Alessandra Dedei, Kinga Dobrowolska, Paola Gagliardi, Noemi Gloor, Alessandra Gregori, Laura Hodel, Rachel Howkins, Patricia Isler, Alexandra Jaszowski, Jasmin Koch, Luca Krenger, Bente Lowin Kropf, Nina Müller, Heike Reimann, Pauline Robert-Charrue, Maja Schärer, Sarah Singh, Fabio Soares, Laura Sopa, Tanja Zepf, Catarina Zweidler.

References

- Baddeley, Alan D. & Graham Hitch. 1974. Working memory. In Gordon H. Bower (ed.), *Psychology of learning and motivation*, 47–89. Cambridge Massachusetts: Academic Press.
- Barcelona European Council. 2002. *Presidency conclusions*.
- Baumert, Jürgen, Johanna Fleckenstein, Michael Leucht, Olaf Köller & Jens Möller. 2020. The long-term proficiency of early, middle, and late starters learning English as a foreign language at school: A narrative review and empirical study. *Language Learning* 70(4). 1091–1135. DOI: 10.1111/lang.12414.
- Berthele, Raphael. 2020. The extraordinary ordinary: Re-engineering multilingualism as a natural category. *Language Learning* 71(S1). 80–120. DOI: 10.1111/lang.12407.

- Berthele, Raphael & Isabelle Udry. 2019. Multilingual boost vs. cognitive abilities: Testing two theories of multilingual language learning in a primary school context. *International Journal of Multilingualism*. 1–20. DOI: 10.1080/14790718.2019.1632315.
- Carroll, John B. 1964. The prediction of success in intensive foreign language training. In *Training research and education*, 87–136. Pittsburgh: University of Pittsburgh Press.
- Carroll, John B. & Stanley M. Sapon. 1959. *Modern language aptitude test*. <http://psycnet.apa.org/psycinfo/1960-03629-000>.
- Christiner, Markus & Susanne M. Reiterer. 2015. A mozart is not Pavarotti: Singers outperform instrumentalists on foreign accent imitation. *Frontiers in Human Neuroscience* 9. DOI: 10.3389/fnhum.2015.00482.
- Dale, Philip S., Nicole Harlaar, Claire M. A Haworth & Robert Plomin. 2010. Two by two: A twin study of second-language acquisition. *Psychological Science* 21(5). 635–640.
- Edelenbos, Peter, Richard Johnstone & Angelika Kubanek. 2006. *The main pedagogical principles underlying the teaching of languages to very young learners*.
- EDK. 2004. *Sprachenunterricht in der obligatorischen Schule: Strategie der EDK und Arbeitsplan für die gesamtschweizerische Koordination*. Bern: Schweizerische Konferenz der Kantonalen Erziehungsdirektoren.
- Enever, Janet. 2011. *ELLiE early language learning in Europe*.
- Erard, Michael. 2012. *Babel no more: The search for the world's most extraordinary language learners*. New York: Free Press.
- Euridyce & Eurostat. 2012. *Key data on teaching languages at school in Europe*.
- European Commission. 2004. *Promoting language learning and linguistic diversity*. Office for Official Publications of the European Communities.
- Flege, James & Ian R.A. Mackay. 2011. What accounts for “age” effects on overall degree of foreign accent? In Magdalena Wrembel, Malgorzata Kul & Katarzyna Dziubalska-Kolaczyk (eds.), *Achievements and perspectives in the acquisition of second language speech*, vol. 2, 62–82. Bern: Peter Lang.
- Garton, Sue, Fiona Copland & Anne Burns. 2011. Investigating global practices in teaching English to young learners. *ELT Research Papers* 11(1). 1–21.
- Giudici, Anja & Sandra Grizelj. 2016. National unity in cultural diversity: How national and linguistic identities affected Swiss language curricula (1914-1961). *Paedagogica Historica* 53(1/2). 137–154. DOI: 10.1080/00309230.2016.1229348.
- Jaekel, Nils, Michael Schurig, Merle Florian & Markus Ritter. 2017. From early starters to late finishers? A longitudinal study of early foreign language learning in school. *Language Learning* 67(3). 631–664. DOI: 10.1111/lang.12242.

- Johnstone, Richard. 2009. An early start: What are the key conditions for generalized success? In Janet Enever, Jayne Moon & Uma Raman (eds.), *Young learner English language policy and implementation: International perspectives*, 31–41. Reading: Garnet Education.
- Muñoz, Carmen. 2006. *Age and rate of foreign language learning*. Bristol: Multilingual Matters.
- Nikolov, Marianne & Benő Csapó. 2010. The relationship between reading skills in early English as a foreign language and Hungarian as a first language. *International Journal of Bilingualism* 14(3). 315–329. DOI: 10.1177/1367006910367854.
- Pfenninger, Simone. 2016. The literacy factor in the optimal age discussion: A five-year longitudinal study. *International Journal of Bilingual Education and Bilingualism* 19(3). 217–234. DOI: 10.1080/13670050.2014.972334.
- Pinker, Steven. 2003. *The blank slate: The modern denial of human nature*. UK: Penguin.
- Reiterer, Susanne M. 2009. Brain and language talent: A synopsis. In Susanne M. Reiterer & Grzegorz Dogil (eds.), *Language talent and brain activity* (Trends in Applied Linguistics 1), 155–191. Berlin: Walter de Gruyter.
- Reiterer, Susanne M. 2019. *Exploring language aptitude: Views from psychology, the language sciences, and cognitive neuroscience* (English Language Education 16). Cham: Springer.
- Rimfeld, Kaili, Philip S. Dale & Robert Plomin. 2015. How specific is second language-learning ability? A twin study exploring the contributions of first language achievement and intelligence to second language achievement. *Translational Psychiatry* 5(9). 638. DOI: 10.1177/0956797610368060.
- Schweizerische Akademie der Geisteswissenschaften. 2015. *Schulischer Fremdsprachenunterricht in der Schweiz: Argumente zur Debatte*.
- Skehan, Peter. 1998. *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, Peter. 2002. Theorising and updating aptitude. In Peter Robinson (ed.), *Individual differences and instructed language learning*, 69–93. Amsterdam/Philadelphia: John Benjamins.
- Stotz, Daniel. 2006. Breaching the peace: Struggles around multilingualism in Switzerland. *Language Policy* 5(3). 247–265. DOI: 10.1007/s10993-006-9025-4.
- Stromswold, Karin. 2001. The heritability of language: A review and meta-analysis of twin, adoption, and linkage studies. *Language* 77(4). 647–723. DOI: 10.1353/lan.2001.0247.
- Wen, Zhisheng (Edward), Peter Skehan, Adriana Biedroń, Shaofeng Li & Richard L. Sparks (eds.). 2019. *Language aptitude: Advancing theory, testing, research*

and practice (Second Language Acquisition and Research Series). New York: Routledge.

Wen, Zhisheng (Edward), Peter Skehan, Richard L. Sparks, Adriana Biedroń & Shaofeng Li. 2019. Researching language aptitude: From prediction to explanation. In Zhisheng (Edward) Wen, Peter Skehan, Adriana Biedroń, Shaofeng Li & Richard L. Sparks (eds.), *Language aptitude: Advancing theory, testing, research and practice* (Second Language Acquisition and Research Series). New York: Routledge.

Chapter 1

Language Aptitude at Primary School (LAPS): Theoretical framework of the project

Isabelle Udry^{a,b}, Raphael Berthele^a & Carina Steiner^c


^aUniversity of Fribourg, Institut de Plurilinguisme ^bZurich University of Teacher Education ^cUniversity of Bern, Center for the Study of Language and Society

This chapter introduces the theoretical framework of the project Language Aptitude at Primary School (LAPS). We considered the impact of a range of individual difference (ID) variables and environmental factors on children's foreign language proficiency. These variables will be discussed in turn, starting with an overview of the language aptitude construct. ID variables pertaining to general cognitive abilities are discussed next, namely intelligence, working memory (WM), creativity, field independence as cognitive style, and metalinguistic awareness. This is followed by an outline of L2 motivation and related constructs to depict the affective dispositions that underlie foreign language learning, i.e. L2 self-concepts, L2 anxiety, and locus of control. Lastly, we discuss the role of environmental factors, such as socioeconomic status, parent education, and teaching paradigm.

1 Introduction

The aim of the project Language Aptitude at Primary School (LAPS) was to explore the impact of a set of individual difference (ID) variables and environmental factors on young learners' developing foreign language proficiency. Of particular interest was how language aptitude, as defined by Carroll (1958), is involved in child learning, a research topic that has only recently started to attract scholarly attention (see §2.2). The project was carried out in two stages. First, we investigated L2 French and L3 English proficiency cross-sectionally (LAPS I $n = 174$).



Isabelle Udry, Raphael Berthele & Carina Steiner. 2021. Language Aptitude at Primary School (LAPS): Theoretical framework of the project. In Raphael Berthele & Isabelle Udry (eds.), *Individual differences in early instructed language learning: The role of language aptitude, cognition, and motivation*, 1–49. Berlin: Language Science Press. DOI: 10.5281/zenodo.5464741 

Second, we recorded children's development of L2 English proficiency over 1.5 years (LAPS II, $n = 637$). The children were aged 10–12 years and learnt two foreign languages in a minimal input setting with 2–3 lessons a week. More details on the study design can be found in Chapter 2, a concise summary of LAPS I and LAPS II is given in the Introduction to the volume.

In the following, we detail the theoretical underpinnings of the ID variables and environmental factors that were considered in the LAPS project.

2 Language aptitude

2.1 Historical overview of language aptitude research and testing

Language aptitude as a construct associated with language acquisition and learning first emerged in the United States in the late 1920s. Learning a second language (L2) as part of tertiary education was encouraged, but little time and money were allocated to foreign language classes. As a consequence, failure rates in these courses were high (Spolsky 1995). Representatives of various educational institutions expressed their concerns and argued for the use of aptitude tests as a way of selecting only suitable candidates for their programs. Calls for prognostic testing became even more pronounced after World War II, when the US army reported an increased demand for staff with good language learning abilities. As a result, aptitude research aiming to develop efficient tests was encouraged and funded by the government (Stansfield & Reed 2004).

John B. Carroll was the first to conceptualize language aptitude. He administered a range of tests assessing relevant abilities for L2 learning to two Air Force groups (total $n = 168$) attending a one-week intensive training course for Mandarin Chinese (Carroll 1958, 1964, 1958). From a factor analysis, Carroll derived four factors associated with successful language learning, which he termed language aptitude:

Phonetic coding ability: The most important component according to Carroll, consisting of the ability to code auditory phonetic material so that it can be recognized, identified and remembered.

Grammatical sensitivity: The ability to recognize the syntactic function of a particular word in a sentence.

Inductive learning ability: The ability to discover grammatical rules independently and without explanation.

Rote learning ability: The ability to establish associations between sound and meaning quickly and efficiently. In other words, the ability to memorize new words rapidly and a sustained capacity for retrieval.

Skehan (1998) later proposed a reduction of Carroll's four dimensions by combining inductive ability and grammatical sensitivity into one subcomponent called *linguistic ability* or *language analytic ability*, while retaining the other two initial components, thus presenting a three-component model. We use this term in Chapter 10 where we discuss the stability of the language analytic aptitude component.

From the assessment tools used to define the components, Carroll & Sapon (1959) selected five tests for the Modern Language Aptitude Test (MLAT) (Table 1).

2.1.1 Newer test batteries

Shortly after Carroll presented his work, Paul Pimsleur published the PLAB (Pimsleur Language Aptitude Battery) for adolescents from grades 7 to 12 (Pimsleur 1966, Pimsleur & Quinn 1971). Its basic structure is similar to the MLAT, but the PLAB differs in including a measure of inductive learning ability and participants' marks from subjects other than languages. Also, Pimsleur regarded motivation as a prerequisite for L2 learning independent of aptitude and dedicated a separate section to it. The PLAB consists of six parts: 1) grade point average in academic areas other than foreign languages, 2) questionnaire on interest in learning a foreign language, 3) vocabulary (word knowledge in L1 English), 4) language analysis (ability to induce rules in an artificial language), 5) sound discrimination (ability to memorize and recognize new phonetic distinctions), and 6) sound-symbol association.

Over time, a shift in research focus occurred, moving from predicting L2 achievement to explaining the underlying mechanisms of language learning. This inspired the development of new test instruments that connected more with current theories on second language acquisition (SLA) and allowed for assessing aptitude differentially in terms of learning stages or learning contexts.

For example, the CANAL-FT (Cognitive Ability for Novelty in Acquisition of Language (Foreign) Test) by Grigorenko et al. (2000) is based on a cognitive theory of knowledge acquisition (p. 392). The CANAL-F theory states that a crucial ability for foreign language acquisition is the ability to cope with novelty and ambiguity. The test therefore simulates naturalistic learning by gradually introducing participants to an artificial language. It assesses specific mechanisms

Table 1: Modern Language Aptitude Test Battery (MLAT) subtests with short description and assessed components (Carroll & Sapon 1959).

Test	Content	Components
MLAT 1 Number Learning	Participants memorize numbers in an artificial language presented to them aurally. During testing, they hear number sequences and write them down.	Phonetic coding ability; Rote learning ability
MLAT 2 Phonetic Script	Non-words must be assigned to phonetic symbols.	Phonetic coding ability
MLAT 3 Spelling Clues	Multiple-choice vocabulary task with a time limit. Participants are asked to find a synonym for a word presented to them in a phonetic transcription.	Phonetic coding ability (Verbal Learning, high correlation on this factor, but the factor was not retained as an aptitude component due to low factor loadings)
MLAT 4 Words in Sentences	The syntactic function of a particular word in a sentence must be recognized and the word with the same function in another sentence must be identified.	Grammatical sensitivity
MLAT 5 Paired Associates	During a trial phase, participants memorize word pairs (artificial language/English) presented in writing. During actual testing, they must assign the English expression to the corresponding artificial word via a multiple-choice task.	Rote learning ability

relevant for language processing, including selective and accidental encoding, selective comparison, selective combination and selective transfer. It is dynamic as it allows for learning during testing.

The MLAT identifies individuals that are likely to progress fast at the beginning of language learning. In contrast, the Hi-LAB (Doughty et al. 2010, Linck et al. 2013) aims to predict high-level attainment in advanced stages of learning. The test includes measures of working memory, associative memory, long-term memory retrieval, implicit learning, processing speed, and auditory perceptual acuity. However, few papers have been published on the validity of the test (Linck et al. 2013), making it difficult to gauge its relevance.

Neither the Hi-LAB nor the CANAL-FT are publicly available and information on content or administration can only be inferred from the literature; the same goes for the DLAB (Defense Language Aptitude Battery, Petersen & Al-Haik 1976) and the VORD (Parry & Child 1990), two other tests mentioned in the literature which are copyrighted by the US government and only administered to its personnel (Robinson 2002). On the other hand, MLAT and PLAB are commercially licensed (although the MLAT for adult learners seems currently only to be sold to government agencies).¹

A freely available test is the LLAMA (Meara et al. 2001), a computer test battery developed, by Paul Meara and his team at the University of Swansea (UK). The LLAMA battery comprises four parts loosely based on the MLAT: Vocabulary learning (LLAMA B), phonemic discrimination (LLAMA D), sound-symbol correspondence (LLAMA E) and inductive ability (LLAMA F). Instructions and tests are administered with pictograms and visual stimuli. Its language-independence makes the test suitable for all participants, regardless of L1 or level of literacy. However, the LLAMA has not been standardized, a disadvantage that is emphasized by the authors themselves (see <http://www.lognostics.co.uk/tools/llama>). Nevertheless, it has been used by numerous research teams and is considered by many to be a reliable tool in aptitude research (Rogers et al. 2017)

2.1.2 Critical views on aptitude testing

The language aptitude components and the MLAT have been derived from empirical data, rather than a specific theory of foreign language learning. The construct is therefore closely linked to the test instruments that measure it. For this reason, language aptitude has been described as “a construct which is, in fact, nothing

¹According to information gathered from the Language Learning and Testing Foundation <https://lltf.net/aptitude-tests/language-aptitude-tests/modern-language-aptitude-test-2/>, last accessed on January 12, 2021.

more or less than what the test measures” (Sáfár & Kormos 2008: 4). Several inconsistencies between the MLAT subtests and the components they target, have added to the controversy over what the test actually stands for (Carpenter 2008). Most notably, the various subtests cannot be assigned clearly to their corresponding aptitude component. For instance, some subtests cover more than one ability (e.g., part 1 “Number learning” assesses both phonetic coding ability and rote memory). Similarly, some components are measured by several tests (rote memory by parts 1 and 5; phonetic coding ability by parts 1, 2 and 3). On the other hand, no test was designed to tap into inductive ability, due to practical reasons of test administration (Carpenter 2008). This component of language aptitude was thus only weakly assessed in part 1 “Number Learning”. The strong yet poorly specified link between Carroll’s aptitude construct and the MLAT test, makes it difficult to build a concise conceptual aptitude framework. Meta-analytical evidence by Li (2016) reveals that commonly used aptitude measures demonstrate differential predictive validities, suggesting that cross-validation of test batteries is called for to determine the extent to which they tap into the same construct. Yet, large scale investigations of aptitude tests are scarce and only few comparative studies exist (for a discussion see Stansfield & Reed 2019).

The MLAT and its derivatives typically rely on discrete-point testing, i.e. they focus on a particular linguistic form which is measured on an item basis. Participants are not given the opportunity to apply language within a context or show their pragmatic skills, an approach that may be more consistent with communicative teaching methods used today (Singleton 2017). The relevance of MLAT-based tests for meaning-focused learning has thus been questioned on several accounts (Krashen 1981, Stansfield 1989, see also Singleton 2017). This is particularly relevant for early instructed language learning and teaching, which is usually based on communication with a focus on fluency over accuracy (for the context of this study see Chapter 2, §2.1). Nevertheless, MLAT-derived tests have been successfully used with young learners (see §2.2.2) and shown explanatory power for their L2 proficiency. Also, as outlined by Stansfield & Reed (2019), several recent studies with adults conducted at US state institutes² (which were reported to adhere to task-based communicative teaching) also indicate that the MLAT remains a sound predictor for L2 proficiency in these learning contexts.

Even though considerable efforts have been made to develop new test batteries, the MLAT remains widely used in the scientific community. Other tests, such as the PLAB, LLAMA, or Hi-LAB have been modelled on it, highlighting how strongly the Carrollian take on language aptitude is still shaping the field. This may be explained by the fact that designing and validating new tools that

²US Defense Language Institute (Winke 2013) and the US Foreign Service Institute (Ehrman 1998).

consider SLA theories and meet the criteria for test quality is challenging. The fact that some test batteries, such as the Hi-LAB, CANAL-F, VORD, or DLAB, are withheld from the public (Ameringer et al. 2019) impedes the scientific community from finding common ground in conceptualizing these measures. The MLAT has been recognized as the foundation of aptitude research. While it is less suited for validating the aptitude construct as outlined previously, its predictive value for L2 proficiency has been repeatedly demonstrated (Li 2016). From this point of view, its continued use appears legitimate.

2.1.3 New conceptions of language aptitude

With a fading interest in the performance-based selection of students prevalent in the early days, explaining the role of various aptitude components for L2 learning and acquisition has become a main concern for researchers (Li 2019). New models have emerged from this explanatory-interactional approach (for a concise overview see Wen et al. 2017).

The *Macro-SLA aptitude model* by Skehan (2002, 2019) maps aptitude components (and corresponding aptitude sub-tests) onto stages of L2 learning. In a recent conceptualization of the model, Skehan (2019) identifies three general (or macro) acquisitional stages organized around 1) handling sound (input processing and segmentation; noticing); 2) handling pattern (identifying; generalizing; and integrating patterns; handling feedback); and 3) automatizing-proceduralizing (avoiding error; automatization; lexicalization). Skehan argues that aptitude components are implicated differently as L2 development progresses. Phonetic coding ability is associated with initial stages of learning when processing auditory input is crucial (handling sound). The remaining two components are more relevant at advanced stages when acquiring complex language structures is important: Language analytic ability helps to recognize and manipulate speech patterns (handling pattern), whereas memory contributes to retaining and retrieving information (automatization). The model comprehensively integrates constructs from language aptitude research and theories of SLA. Nevertheless, this integration still remains conceptual to a large extent and more empirical support is needed to validate it.

The *Aptitude complexes framework* was conceived by Robinson (2001; see also Robinson 2002) to be applied to instructed foreign language learning. The framework postulates aptitude clusters that consist of cognitive resources (memory, attention, basic processing speed), language-specific abilities (e.g., noticing the gap, memory for contingent speech) and domain-general, primary cognitive abilities that support language acquisition (e.g., perceptual speed or pattern recognition).

Robinson argues that individual learner characteristics reflected in these aptitude complexes are compatible with specific teaching methods. For instance, the aptitude cluster for incidental learning (via oral content) combines well with a communicative classroom setting where linguistic phenomena are mediated implicitly. The practical aim of this framework is to enhance L2 learning by matching teaching method to aptitude complex.

Other models that re-conceptualize language aptitude include the linguistic coding differences hypothesis (LCDH) by Sparks & Ganschow (1991), which takes on the view of learning difficulties and the L1–L2 connection, the distinction between an explicit and implicit language aptitude (Grañena 2012, 2016) or the brain-network-based view on language aptitude in the neuro-scientific perspective (Golestani et al. 2011, Reiterer et al. 2013). Even other models are linked to the development of new test batteries and have been touched upon in §2.1.1: the high level language aptitude battery (Hi-LAB) model with a focus on exceptional language learners and the CANAL-F theory that highlights the ability to deal with novelty and ambiguity in language learning.

2.2 Aptitude differences in children

Due to a focus on student selection for state-funded language programs, early aptitude research was mainly concerned with adults and adolescents. It was not until 1976 that Carroll and Sapon adapted their MLAT (Carroll & Sapon 1959) to create the first test battery for children, the Modern Language Aptitude Test – Elementary (MLAT-E). Still widely used today, it is designed for L1 English speakers between 9 and 12 years of age (grades 3 to 6) and consists of four subtests outlined in Table 2.

The lack of interest in young learners was further owed to the assumption that language aptitude accounts for L2 achievement in adults, but not children (Li 2018). This claim is made with reference to the fundamental difference hypothesis (FDH, Bley-Vroman 1989) and the critical period hypothesis (CPH) popularized by Lenneberg (1967). The FDH and CPH argue that children draw on implicit, domain-specific mechanisms to learn languages. Due to maturational changes, they lose access to the domain-specific abilities upon entering puberty and start to rely on domain-general abilities instead. It is further argued that language cannot be learnt fully by domain-general mechanisms, particularly in relation to grammar and phonology. As will be discussed in the next section, exceptional cases of high attainment in late starters, i.e. individuals that started learning a L2 after completion of the supposed critical period, have therefore been linked by some scholars to above-average levels of language aptitude (DeKeyser 2000), particularly its verbal analysis component.

Table 2: Modern Language Aptitude Test Battery – Elementary (MLAT-E) subtests with short description and assessed components (Carroll & Sapon 1976).

Name of the test	Content	Component	Based on MLAT subtest
1. Hidden Words	Multiple-choice vocabulary task with a time limit. Participants are asked to find the L1 English definition for a word presented to them in a phonetic transcription.	English vocabulary Phonetic coding ability	Spelling Clues
2. Matching Words	Recognizing the function of a particular word in a sentence.	Grammatical sensitivity	Words in Sentences
3. Finding Rhymes	In a multiple-choice task, participants need to select words that rhyme.	Phonetic coding ability (ability to hear speech sounds)	Not in the original MLAT
4. Number Learning	Participants memorize numbers in an artificial language presented to them aurally. During testing, they hear number sequences and write them down.	Phonetic coding ^a Rote memory	Number Learning

^a As stated in the MLAT-E manual (Carroll & Sapon 1976: 2) the Number Learning subtest also taps into what the authors refer to as “a special ‘auditory alertness’ factor which would play a role in auditory comprehension of a foreign language” However, “auditory alertness” was not retained as a subcomponent of Carroll’s aptitude construct.

2.2.1 Child aptitude and ultimate L2 attainment

One perspective on child aptitude is its effect on ultimate L2 attainment in adults. Based on work by Johnson & Newport (1989), DeKeyser (2000) examined the role of language aptitude (along with age of arrival and years of schooling) as predictors for L2 English grammaticality judgment (GJ) accuracy among 57 Hungarian immigrants to the US. The participants were divided into groups of early ($n = 15$) and late arrivals ($n = 42$), as well as high aptitude ($n = 15$)³ and average- or low-aptitude ($n = 42$). Only few late arrivals reached scores within the range of early arrivals on the GJ test. Those who did all had high levels of language aptitude, operationalized as verbal analytical ability. Overall, language aptitude was not predictive of GJ accuracy. However, for late arrivals, GJ scores were significantly and positively correlated with verbal analytical ability. From this, the author concluded that language aptitude plays a role for ultimate attainment only for late starters, thus providing an explanation for those highly successful individuals that challenge the CPH.⁴ Similarly, in a study with 65 Chinese learners of Spanish, Grañena & Long (2012) found aptitude effects only for late learners whose first contact with the L2 happened between the ages of 16 and 29 years ($n = 18$). Significant correlations between aptitude and pronunciation, aptitude and lexis and aptitude and knowledge of collocations were found, but not between aptitude and morphosyntax.

Abrahamsson & Hyltenstam (2008) provided evidence on the role of language aptitude for late starters ($n = 11$) with 42 near-native L2 speakers of Swedish with L1 Spanish. But contrary to DeKeyser (2000), the authors also found aptitude effects for early starters. Yet, the authors concluded that finding a few individuals with high aptitude “does not justify a rejection of the critical period hypothesis” (Abrahamsson & Hyltenstam 2008: 503).

Adding to the mixed findings is Grañena (2012), who examined age and aptitude in relation to ultimate L2 attainment with 100 Chinese-Spanish bilinguals. She identified two types of language aptitude: One for explicit learning (termed analytic ability) and one for implicit learning (defined as sequence learning ability) and found that both affected early L2 learners’ attainment.

Several things may contribute to the inconclusiveness of these results. First, proficiency and aptitude were operationalized differently and therefore measured with different tools, making it difficult to compare findings. For instance,

³The 15 participants in the high aptitude group are not identical to the $n = 15$ of the early arrival group.

⁴See also Vanhove 2013 for a critical view on what counts as statistical evidence in favor or against CPH.

aptitude was assessed with the LLAMA (Meara et al. 2001) by Abrahamsson & Hyltenstam (2008), and Grañena & Long (2012) while DeKeyser (2000) used a sub-test of language analysis from a Hungarian aptitude test (adapted from the MLAT Words in Sentences subtests by Ottó 1996). L2 proficiency was measured by an aural GJ task in DeKeyser (2000), an aural and written GJ task in Abrahamsson & Hyltenstam (2008), or several tests of different language domains, including pronunciation, lexis, and morphosyntax by Grañena & Long (2012) and Grañena (2012). Furthermore, different criteria were applied to define age groups: With cut off points for early learners at 12 years (Abrahamsson & Hyltenstam 2008) or 16 years (DeKeyser 2000), Grañena & Long (2012) and Grañena (2012) had three groups with ages of onset between 3–6, 7–15 and 16–29.

These studies were concerned with aptitude effects on ultimate L2 attainment in naturalistic contexts. Despite similar L2 learning conditions, participants may still have experienced very diverse linguistic environments including some form of formal instruction. A variety of variables, beyond aptitude or age of onset, may therefore account for ultimate achievement. As pointed out by Birdsong (2014), the DeKeyser study was built around critical period effects in relation to age of arrival and L2 proficiency. As a result, the explanatory power of education (assessed as years of schooling) was not fully explored. Reanalyzing the same data, Birdsong (2014) found that years of schooling was in fact the most robust predictor of grammatical proficiency with significant correlations in all age and aptitude groups. Education and aptitude, however, did not correlate for any age group, indicating that the two variables make independent contributions.

2.2.2 Studies with children

With early instructed language learning being introduced across Europe (see Introduction, §3), the age factor has gained in importance on the research agenda and has led to the publication of several studies with children. They are concerned with 1) evaluating the predictive power of language aptitude (and its sub-components) for L2 proficiency, 2) developing test batteries for young learners, 3) the stability of language aptitude, and 4) its relationship with other constructs, such as metalinguistic awareness or motivation. Some studies combined these aspects, for instance, validation studies of newly developed aptitude tests by Kiss & Nikolov (2005) or Suárez Vilagran (2010) also investigated age-related questions. The most notable findings will be presented in the following.

First of all, it is worth pointing out that despite assumptions drawn from FDH and CPH that aptitude may be irrelevant for child learning, studies have consistently found language aptitude to be a predictor of L2 proficiency in young

learners (Bialystok & Fröhlich 1978, Kiss & Nikolov 2005, Kiss 2009, Suárez Vilagran 2010, Muñoz 2014, Tellier & Roehr-Brackin 2017, Roehr-Brackin & Tellier 2019).

For instance, Tellier & Roehr-Brackin (2017) tested 178 8- to 9-year-old English-speaking beginning learners of French on metalinguistic awareness and language aptitude (tested with a British version of the MLAT-E). Language aptitude was shown to have a significant effect on children's progress in L2 French classes with a form-focused element.

Kiss & Nikolov (2005) developed, piloted and validated an aptitude test in Hungarian, modelled on the MENYÉT (Ottó 1996, in Kiss & Nikolov 2005), a Hungarian adaptation of the MLAT (Carroll & Sapon 1959) and the PLAB (Pimsleur 1966). Their final version for young learners consists of 4 subtests (targeted aptitude component in brackets):

1. *Hidden sounds*: Associating sounds with written symbols (phonetic coding);
2. *Words in sentences*: Identifying semantic and syntactic functions in Hungarian sentences (grammatical sensitivity);
3. *Language analysis*: Recognizing structural patterns in an artificial language, based on part 4 of the PLAB (inductive ability);
4. *Vocabulary learning*: Paring words and phrases in an artificial language with Hungarian equivalents (rote memory).

Kiss & Nikolov (2005) administered the aptitude test along with measures of motivation and English proficiency (listening, reading, writing) to 419 12-year-old children learning English as a foreign language. Time of exposure to English at school and in private tuition ranged considerably from 100 to 1,085 hours ($M = 343$; $SD = 131$). Multiple regression analysis indicated that language aptitude was the best predictor of outcomes, explaining over 20% of the variance in L2 English proficiency. Motivation also made a significant contribution, explaining 8% of the variance. Moreover, the authors found a weak correlation between time spent on learning and aptitude scores. From this they concluded that language aptitude in the Carrollian sense did not improve with “the amount of time used for practice and exposure” (Kiss & Nikolov 2005: 134).

Kiss (2009) adapted and piloted a version of this Hungarian test battery for 8-year olds. This was done with a practical aim in mind, i.e. selecting 26 children out of 52 for a dual Hungarian-English language program. After one year of study in

the bilingual class, the children ($n = 25^5$) were tested for English proficiency with a short interview. Their progress was also rated by their teachers. Achievement was related to the aptitude scores taken before they had entered the program. Most notably, the author compared the results from the 8-year-olds to those from 12-year-olds from a previous study. She found that the 12-year-olds performed much better on the vocabulary learning subtest than the younger children. Kiss (2009) argued that the older children had more language learning experience and better developed strategies. Based on the idea that aptitude malleability can be evidenced by increased group averages, the author concluded that language aptitude is dynamic and shaped by language experience, at least up to the age of 12.

Suárez Vilagran (2010) validated adaptations of the MLAT-E into Spanish (MLAT-ES, Stansfield & Reed 2005) and Catalan (MLAT-EC, Suárez Vilagran 2010) with 629 Spanish-Catalan bilingual learners of English from grades 3 to 7 (aged 8,3–14,9). MLAT-ES and MLAT-EC are structured like the MLAT-E and, unlike the Hungarian version, they do not contain a test for inductive ability. There are four tasks (targeted aptitude component in brackets):

1. *Hidden words*: Ability to utilize previously learned sound-symbol associations (phonetic coding);
2. *Matching words*: Sensitivity to grammatical structures presented in the target language (grammatical sensitivity);
3. *Finding rhymes*: Ability to recognize sequences of orthographically presented speech sounds (phonetic coding);
4. *Number learning*: Memorizing numbers in an artificial language (rote memory and phonetic coding).

Suárez Vilagran (2010) measured foreign language proficiency with a multiple-choice listening test and a cloze passage in all grades. In addition, children in grades 5, 6, and 7 took a dictation test. The author found both test batteries to be valid measures for predicting L2 proficiency, although not for speaking. In terms of the subcomponents, the Hidden Words test (phonetic coding) showed the lowest correlations with proficiency across all grades, while Matching Words (grammatical sensitivity) and Finding Rhymes (ability to hear speech sounds) were significantly correlated with L2 proficiency from grades 4 to 7. The author also

⁵One child was absent on the day of testing.

highlighted some age-related findings: Overall, mean scores stabilized between grades 6 and 7 (Suárez Vilagran 2010: 349). Grade 3 showed notable patterns in several respects: Correlations between aptitude and language proficiency for grade 3 were consistently lower than for other grades. Similar to arguments put forward by Kiss (2009), the author relates this to cognitive development, notably less developed strategies for problem-solving, for encoding and memorizing information. Also, grade 3 students scored lower on metalinguistic awareness tests than the older participants.

2.2.3 Child aptitude and memory

Moreover, Suárez Vilagran (2010) found that correlations between the Number Learning test (rote memory) and L2 proficiency decreased as children got older, suggesting that memory is more important for younger learners than for older ones. This finding is in line with the high importance of exemplar-based learning in child L1 acquisition as argued, e.g., by Tomasello (2005). In such a usage-based framework, variations in memory capacity are expected to be strongly associated with language learning particularly in young learners.

Investigating the relationship between language aptitude components and L2 proficiency, Muñoz (2014) also found slightly stronger effects for rote memory on language outcomes in 48 Spanish-Catalan bilinguals aged 10–12 years, learning L2 English. The author administered the MLAT-ES along with measures of L2 listening, reading, writing, and speaking. Her results corroborate findings, such as the ones from Suárez Vilagran (2010), that “children rely on memory to a large extent”, (Muñoz 2014: 64). While Muñoz (2014) highlights children’s reliance on memory, she also emphasizes the importance of the other aptitude components; in particular the author suggests that language-analytic abilities are likely to be the key component for high achievement.

Memory did not always yield the strongest correlations with L2 proficiency: Kiss & Nikolov (2005: 140) found both memory and analytical abilities to be relevant for L2 proficiency and in Roehr-Brackin & Tellier (2019) analytic ability emerged as the strongest predictor, followed by a measure of phonetic ability.

2.2.4 Aptitude in children with beginning literacy skills

The studies outlined so far have found aptitude effects for children at the primary and early secondary level. Alexiou (2009; see also Milton & Alexiou 2006) was interested in even younger children with beginning or no literacy skills. Based on the work by Esser & Kossling (1986), the author designed the YLAT (Young

Learners Aptitude Test) for children between 5 and 7 years. It contains tasks that partially overlap with the Carrollian aptitude components. For instance, inductive ability is assessed with a task in which colors represent groups of objects (blue for flowers, white for animals, etc.) that the child must discover and systematize. Long-term memory is tested by an adapted version of the MLAT subtest Paired Associates, which contains only visual stimuli. Short-term memory, semantic integration, spatial skills, and reasoning ability (sequencing narrative elements) are also assessed by the YLAT. In a study with Greek learners of English aged 5–7 years ($n = 191$), Alexiou (2009) found significant correlations ranging from 0.33 to 0.65 between the different dimensions assessed by the subtests and L2 vocabulary (receptive and productive). Her findings corroborate observations that aptitude explains individual differences in child L2 learning from an early age.

2.2.5 Child aptitude and musical talent

Indications for a link between musical talent and foreign language learning stem from Christiner & Reiterer (2018: $n = 35$) and Christiner (2018: $n = 36$) who investigated pre-schoolers' musical ability and speech imitation ability as an aspect of language aptitude. Children between 5 and 6 years of age were tested for their ability to discriminate paired musical statements, singing ability, ability to remember strings of numbers and ability to repeat Turkish, which was an unfamiliar language to them. Participants with good performance on the musicality measure also scored high on the imitation tasks and had high working memory capacity compared to participants with lower scores on the musicality test. The authors concluded that musical talent and speech imitation aptitude are related in children.

2.3 Aptitude stability

Whether language aptitude is a stable trait or an ability susceptible to treatment is an ongoing debate in aptitude research. If aptitude was stable (and possibly also innate), success or failure at language learning would be largely predetermined. If, on the other hand, aptitude was trainable, it could be used to enhance foreign language instruction. The question also deserves attention in relation to children who are still developing mentally and physically in various ways.

In traditional models, aptitude was assumed to be a stable trait (Skehan 1998, Singleton 2017). A long-term study by Skehan (1986, Skehan & Ducroquet 1988) has been widely held to corroborate this view. The authors assessed the language

development of children in their L1 and 13 years later in their L2. Some measures in the L1 collected between the ages of 39 months and 57 months proved to be related to measures of L2 acquisition. In particular, L1 vocabulary and early mean length of utterance were correlated with later L2 aptitude test scores. From this link between L1 acquisition indices and L2 learning, the authors concluded that an aptitude for language learning is a stable individual characteristic. In Chapter 9, the development of L1 German reading comprehension and L2 English proficiency also shows similar predictive variables.

Early research on aptitude with its objective of selecting the apt individuals, explicitly or implicitly assumed individual differences in aptitude to be innate (Carroll 1964: 122, 1973: 8). More recent research on the genetic contribution to L2 learning seems to support the idea that substantial proportions of the variability in learning outcomes are explained by genetics (Stromswold 2001, Rimfeld et al. 2015, Plomin 2019). Recent scholarly approaches in aptitude research have attempted to explain processes of SLA, rather than predict learning outcomes (Wen et al. 2019). As a result, some authors now model language aptitude as an array of abilities that can potentially be developed. For instance, Grigorenko et al. (2000: 401) refer to a form of “developing expertise rather than an entity fixed at birth”. Carroll later expressed himself neutral on the issue of aptitude stability, arguing that no empirical evidence was available to decide on the matter (Carroll 1981: 86). In the same paper, Carroll (1981: 84) suggested that his initial aptitude components could be modelled as “more or less enduring characteristics” and as a “current state”.

Empirical studies of construct stability are scarce and generally seem to confirm its malleability (Sáfár & Kormos 2008, Suárez Vilagran 2010, Roehr-Brackin & Tellier 2019). It is worth noting that, except for Roehr-Brackin & Tellier 2019, researchers relied on cross-sectional data to infer developmental patterns, rather than multiple indications from the same participants collected longitudinally. Moreover, the designs were based on the premise that language experience (i.e. instructed learning), leads to developments in language aptitude, especially the language analysis component (grammatical sensitivity and inductive ability). Gain scores in aptitude measures were therefore interpreted as an indication of aptitude development. However, these gain scores may be due to other cognitive changes, rather than changes in language-analytic ability. Children in particular are still evolving in terms of literacy and reasoning skills. Due to general developmental processes, young learners are expected to do better at the same aptitude test as they mature. An increase in aptitude mean scores with growing age may not be a reliable indicator for aptitude malleability. In order to ascertain if other developmental mechanisms are implicated in improved test results, these results

would need to be compared to age-normed charts, such as provided for instance for standardized intelligence tests. These charts allow for classifying an individual's score in comparison to a representative sample from the same age group. As we argue in Chapter 10, another way of investigating aptitude stability is to look for individual patterns of development in longitudinal data with several measurements for the same participants.

2.4 Language aptitude and pedagogy

2.4.1 Aptitude Treatment Interaction (ATI)

The description of different aptitude complexes outlined in §2.1.3 (Robinson 2001, 2002) opens up new perspectives for researching and planning foreign language teaching: If learners have different strengths, it is to be expected that successful learning depends on the way these individual strengths can be attended to in the classroom. The assumption that matching aptitude profiles with certain teaching methods will increase learning gains is at the core of the aptitude-treatment-interaction (ATI) approach.

Based on founding work by Snow (1991), Robinson extended the ATI framework to L2 learning and teaching. To date, ATI has explored the presumed interface between aptitude and learning environment along the lines of: 1) Implicit and explicit instruction, which both seem to be influenced by IDs in aptitude (de Graaff 1997, Robinson 1997, Williams 1999); 2) deductive and inductive instruction, with current results suggesting that a deductive approach combined with extensive opportunities for production seems to benefit all learner types, regardless of aptitude profiles (Erlam 2005); and 3) corrective feedback. A synthetic review by Li (2017) revealed that language aptitude was moderately correlated with the effectiveness of corrective feedback ($r = 0.42$), and more strongly with explicit feedback ($r = 0.59$) than implicit feedback ($r = 0.32$).

So far, only one study has explored the connection between aptitude profiles and instructional treatments on a large scale. Wesche (1981) derived aptitude profiles for each participant from three different sources: Aptitude tests (MLAT and PLAB), L1 proficiency measures and an interview with an experienced teacher. Pairs of learners with the same profile were assigned to different instructional groups: One person was taught according to their profile, the other one according to a method that was unsuitable for their profile. The choice was between three teaching methods: The analytical approach (best suited for highly analytical students with strong L1 skills and perfectionist tendencies); the functional approach (appropriate for students with a relatively restricted command of their L1, yet

with good memory capacity); and the audio-visual method (the most common way of teaching at the time of the study and best-suited for non-type-specific learners). After 55 lessons, participants who were exposed to a suitable teaching method achieved higher L2 proficiency scores and reported more pleasure in language learning than their counterparts.

2.4.2 Aptitude and classroom practice

Several contributions from ATI to the foreign language classroom are worth contemplating (Cook 2001, Ranta 2008).

In the prognostic view, aptitude tests are used to make inferences about students' development, a practice that reminds us of the early days of aptitude testing. By stipulating certain thresholds of scores, students can be selected or dispensed from language classes, depending on how well they reach the prescribed levels. Aptitude tests have also been used for student placement, with scores being interpreted as an indication of how well an individual will be likely to cope with foreign language instruction. Remember that aptitude tests are reliable predictors for L2 outcomes. They provide information on cognitive-linguistic aspects of the individual but say little about, for instance, motivation to learn the language. In order to fully gauge a student's potential, it is advisable to supplement aptitude tests with assessments of motivation, general learning abilities, and careful consideration of the implications for the student's academic future.

Findings from the exploratory-interactional approach are suited for diagnostic assessment purposes, i.e. for counselling students based on their aptitude strengths and weaknesses. For example, students with good language analytic ability could be advised to choose explicit learning. Memory-oriented students, on the other hand, could be guided towards communicative classes, since they are likely to learn through modelling (see section 3.2). However, this implies that schools can actually provide an infrastructure that accommodates these different choices. A tangible example of how aptitude clusters could be used for counselling seems to come from Doughty (2013): Students' scores from the Hi-LAB (see 1.1.1) are visualized in a so-called aptitude profile card, which is available to learners and teachers along with advice for individual learning. Unfortunately, there is little information available on the effectiveness and exact implementation of these cards.

Wesche's (1981) intervention study discussed in §2.4.1 is the only large-scale attempt to assign entire groups of students to a type of instruction based on their aptitude profiles for a longer period of time. Her study took place in a particular educational context with adults when learning was mainly form-focused and

communicative teaching was the alternative option. Current teaching practices and learning settings differ quite considerably, especially for children. An alternative to Wesche's approach consists in using different instructional techniques simultaneously in the same classroom, adapting continuously to individual learner requirements. For instance, if high-aptitude students benefit more from explicit corrective feedback and low-aptitude learners from implicit corrective feedback (Li 2017), then both types should be used by the teacher during a lesson based on students' needs.

Also, drawing on language aptitude for internal differentiation regarding treatment within clusters (groups of learners, classes) assumes that there is indeed an interaction between aptitude and instruction. Erlam (2005) investigated such an interaction with three teaching styles (inductive, deductive and structured input) in relation to the aptitude profiles of 60 Anglophone learners of French at secondary school. The author found that a deductive approach combined with extensive opportunities for productive output was beneficial to all learners, regardless of their aptitude profile. Her results suggest that a particular type of instruction (i.e. deductive + productive output) may diminish the influence of individual aptitude differences. It would therefore suffice to teach according to this method without providing aptitude-based differentiation. The kind of finding reported by Erlam (2005) is worth pursuing as it may offer opportunities for more efficient lesson planning.

One last line of application worth mentioning is linked to the potential trainability of language aptitude, namely language analytic ability, suggested by some authors (Grigorenko et al. 2000, Sáfár & Kormos 2008). Fostering these abilities is expected to positively affect L2 learning. To date, however, the direct effects of such a training on L2 proficiency remain to be clearly ascertained empirically for primary school children.

Whereas it is uncontested that learners vary in terms of their aptitude to learn new languages, the practical consequences of this insight for the foreign language classroom are not obvious. In the previous section, we have presented some feasible suggestions which are nonetheless rarely implemented at schools today. Moreover, little is said in the literature on how to connect empirical findings from ATI to classroom practice. This may be due to several reasons. Conducting ATI research is indeed challenging given the wide range of factors that affect language learning, i.e. type of instruction, cognitive processes, IDs. Due to this complexity, ATI studies are usually carried out over short periods of time and with small samples. Because few studies have been conducted within the ATI line of research, too little is known about the interaction between aptitude and treatment. Moreover, the educational relevance of individual learning styles

in general – indeed, their very existence – remains highly contested in the field of educational psychology (Riener & Willingham 2010). The results discussed in §2.4.1 seem promising but we believe that many more similar studies would be required to make sound claims about the effects of ATI based learning settings and to counter the well-argued objections to learning style claims in the literature (Pashler et al. 2008, Riener & Willingham 2010).

3 General cognitive abilities or general learning abilities

In order to explore the interplay between domain-specific and domain-general abilities, we included ID variables pertaining to what we refer to as general cognitive abilities or general learning abilities. We start by clarifying some aspects of their relationship with language aptitude. Next, we outline the constructs underlying intelligence, working memory (WM), creativity, and field independence. Finally, we discuss metalinguistic awareness and the language analysis subcomponent of aptitude which are hypothesized to be closely linked.

3.1 General cognitive abilities and language aptitude

Carroll (1964: 89) described language aptitude as “a fairly specialized talent (or group of talents), relatively independent of those traits ordinarily included under ‘intelligence’”. His statement was underpinned by the observation that intelligence tests were quite unsuccessful in screening individuals for successful language learning (Carroll 1964). Currently, general psychological mechanisms and processes are often highlighted as underlying language learning and acquisition. Nevertheless, aptitude test items and instructions are usually mediated by language, so the construct is at least language-related. Based on these observations, Skehan (2019) has recently argued for a complementary view, suggesting that domain-general and domain-specific capacities co-exist and should be equally reflected in aptitude research.

Recent scholarly work has often adopted a domain general perspective, investigating aptitude and intelligence (Grañena 2012, 2013), the role of different memory systems (declarative, procedural, Carpenter 2008, Morgan-Short et al. 2014), or working memory as a distinct aptitude component (Wen 2019). Also, new test batteries include general cognitive measures (i.e. working memory and processing speed in the Hi-LAB, Linck et al. 2013, see §2.1.1). The connection between language aptitude and general cognitive abilities is likely to remain an important research focus in the future.

3.2 Intelligence

3.2.1 Definitions and operationalizations of intelligence

The earliest model of intelligence goes back to Charles Spearman (1904) who proposed a two-factor model with a general factor (g) plus other, more specific abilities (s). The g factor is thought of as general mental ability involving more or less complex mental activities, such as recognition, recall, speed, visual-motor abilities, motor abilities, reasoning, comprehension and hypothesis-testing activities (Sattler 2001). Several other hierarchical models were derived from Spearman's work.⁶ More recently, non-hierarchical models have been put forward which conceive of different forms of intelligence as existing independently and equivalently of each other. Widely known is Gardner's (1983) theory of multiple intelligences, which includes social-emotional, musical, physical-kinaesthetic, interpersonal and intrapersonal forms of intelligence. Sternberg (1985, 2002) theorizes three forms: Analytical, creative and practical intelligence, which are drawn on "to adapt to, shape, and select environments" (Sternberg 2002: 15).

For the present study, intelligence was operationalized according to Cattell's (1943) two component theory which postulates a type of fluid and crystallized intelligence. Fluid intelligence refers to a general ability to think and problem solve, largely independent of cultural influences. Fluid intelligence is considered an important prerequisite for acquiring new information and therefore learning. Cattell argues that fluid intelligence is largely fixed at birth. In contrast, crystallized intelligence consists of knowledge and skills acquired throughout life. It increases with growing experience and is thought to be influenced by culture and language ability. The two develop differently over the life span, with crystallized intelligence increasing over the years until it stagnates at some point, and fluid intelligence decreasing with age. The two are considered separate factors linked by a common overarching factor g. Cattell's model has left its mark on intelligence testing with the development of so-called Culture Fair Tests. These tools are designed to tap into fluid intelligence, thus cancelling out cultural differences that may affect performance. In our study we used the CFT 20-R which is a standardized version for German speaking children from the age of 8 (see also Chapter 2).

⁶For instance, Thorndike's multifactor theory of intelligence in the late 1920s, Thurstone's multidimensional theory of intelligence in the 1930s or Vernon's hierarchical theory of intelligence in the 1950s (Sattler 2001).

3.2.2 Intelligence and foreign language learning

Early studies that dealt with the relationship between intelligence and L2 acquisition reported high correlations between the two (Spolsky 1995: 327f). In contrast, later work emphasized two independent constructs (Gardner & Lambert 1965, Skehan 1986). Recently, a more differentiated view considering the interaction between various subcomponents of aptitude and intelligence has emerged. Sasaki (1996) assessed L2 English proficiency and aptitude⁷ as well as two measures of general intelligence (verbal and reasoning) in Japanese students. The study indicated correlations between intelligence and language analytic abilities, although phonetic coding ability and rote memory (as defined by Carroll) correlated only weakly with measures of general intelligence.

In two studies with different samples (100 adult Chinese-Spanish bilinguals and 186 adults with different L1s), Grañena (2012, 2013) found intelligence to be associated with explicit learning. The author administered a comprehensive test battery comprising what she refers to as explicit aptitude (LLAMA B, E, F, Meara et al. 2001), implicit aptitude (LLAMA D and a probabilistic serial reaction time task), and intelligence (according to the author, with a test corresponding roughly to an assessment of fluid intelligence).⁸ Statistical analysis confirmed the presence of two distinct aptitude dimensions associated with explicit and implicit L2 learning mechanisms. General intelligence correlated strongly with the former, explicit factor.

Wesche et al. (1982) concluded that aptitude (measured with the MLAT) and intelligence (Primary Mental Abilities Test PMA assessing reasoning ability, word fluency, verbal comprehension, facility with numbers, spatial visualization, and rote memory)⁹ are relatively distinct factors, but they are not independent of one another. These findings were interpreted in a hierarchical model subsuming specific abilities important to instructed language learning under a more encompassing general ability or under general intelligence as postulated in Spearman's *g* factor.

Li (2016) explored the construct validity of language aptitude in a meta-analysis including 66 studies with 109 unique samples and 13,035 foreign language learners. The author found a strong correlation ($r = 0.64$) between aptitude and intelligence. This may be due to similarities between measures of aptitude and intelligence. For instance, both usually include tests of L1 vocabulary and memory. The reported correlation is not strong enough to speak of an identical construct

⁷JLAB (Japanese Language Aptitude Battery) based on the MLAT.

⁸Spanish version of the General Ability Measure for Adults (GAMA).

⁹Assessed with the PMA Primary Mental Abilities Test.

(Li 2016). Nevertheless, the author argues for further examining this overlap in order to clarify construct validity. Indeed, if language aptitude is not distinguishable from abilities required in other areas of academic learning, its existence as a construct becomes redundant.

3.3 Working memory (WM)

Working memory (WM) is associated with the ability to temporarily store and manipulate information and thus underpins our capacity for complex cognitive behavior (Baddeley 2003). A widely held model for explaining language acquisition and processing is the multi-component model of WM by Baddeley & Hitch (1974, Baddeley 2000). It consists of the central executive that acts as an attentional control system for the flow of information. The central executive is supported by three slave systems: 1) The phonological loop (also referred to as phonological short-term memory), which processes verbal and acoustic information, 2) the visuo-spatial sketchpad, which deals with visual information and 3) the episodic buffer, which integrates and temporarily stores information from the different modalities. Several studies have corroborated the presence of Baddeley's WM structure in children (for an overview see Boyle et al. 2013).

The phonological loop contains two further subparts, 1) a short-term phonological store where memory traces of auditory information are held for a few seconds and 2) an articulatory rehearsal component that keeps information activated to prevent time-based decay. Developmental studies suggest that the phonological store is established by the age of 3 with the capacity for subvocal rehearsal emerging around the age of 7 and increasing into adolescence (Hasselhorn & Grube 2003). The phonological loop has been described as central to L1 vocabulary acquisition and the development of spoken language in general (Baddeley et al. 1998). It has also been linked to L2 development in children and adults, more specifically in learning new sound patterns (Speciale et al. 2004), L2 grammar (French & O'Brien 2008) and L2 oral performance (O'Brien et al. 2006).

The visuo-spatial sketchpad is formed by the age of 4 and no further significant developmental changes seem to occur in this subsystem between 5 and 10 years (Hasselhorn & Grube 2003). It has been associated in particular with learning spatial routes and faces and may be implicated in the acquisition of arithmetic skills (Gathercole & Pickering 2000: 179).

The episodic buffer was later added to Baddeley's WM model to account for language performance in individuals with impaired phonological memory (Baddeley 2000). Despite deficiencies in the phonological loop, these individuals were able to perform tasks that involve processing of complex auditory and visual

information, such as recalling narratives or remembering sets of playing cards dealt in a game. The episodic buffer was proposed as a possible explanation: It is hypothesized as a system which is able to integrate information from all sub-systems and from long-term memory into a unitary episodic representation (Baddeley 2000: 417). The episodic buffer is conceived of as an interface between the other WM components and long-term memory.

In relation to children, there is ample evidence associating WM performance to complex cognitive abilities which are likely to influence academic achievement (Gathercole & Pickering 2000: 175). More specifically, the central executive has been linked to vocabulary acquisition, reading and arithmetic skills. The phonological loop is particularly related to language acquisition, i.e. long-term learning of the sound patterns of new words (Gathercole & Pickering 2000). WM capacity in all the components mentioned is limited and has been shown to increase throughout childhood until the individual reaches young adulthood (Hasselhorn & Grube 2003).

3.3.1 Measuring WM capacity

Measures of WM capacity distinguish between 1) the storage and processing functions and 2) the verbal (domain specific) and non-verbal (domain general) dimension (Linck et al. 2014, Wen 2015). Simple span tasks assess storage only, i.e. they are indicative of short-term memory. Simple tasks include word and digit span tests that require participants to recall increasing numbers of unrelated words or numbers (Juffs & Harrington 2011). For instance, the forward digit span is a non-verbal simple test in which increasing numbers of random digits are presented until the individual reaches maximum recall capacity. Complex span tasks assess both storage and processing, i.e. they pertain to executive WM (EWM). A frequently used measure of complex WM is the Reading Span task (RST, Daneman & Carpenter 1980) in which individuals need to simultaneously read aloud and comprehend sentences and recall the final word of each sentence. The Listening Span task is equivalent to the RST and assesses auditory storage and processing. A non-verbal option for a complex task is the Operation Span test (Turner & Engle 1989) in which sentences are replaced with simple arithmetic equations. The Backward Digit Span Task (BDS) also reduces the language load (Kormos & Sáfár 2008). In the BDS, participants are presented with an increasing number of random digits which they have to recall in reverse order (Juffs & Harrington 2011).

Validity and reliability of widely used measures of WM capacity (including counting span, operation span, and reading span) have been documented by Con-

way et al. (2005) in a methodological review. A meta-analysis conducted by Linck et al. (2014: 861) on WM and L2 comprehension and production¹⁰ revealed that complex span tasks are more predictive of L2 outcomes than simple span tasks, indicating that EWM may be more strongly implicated in L2 use than short-term memory. According to Juffs & Harrington (2011: 158) the RST and the Listening Span test are particularly successful predictors of L2 learning.

3.3.2 WM and language aptitude

Carroll's rote memory component stems from an associative representation of memory serving as a static and passive storage space for information. Early aptitude tests usually measured rote memory with word lists that required individuals to map unknown words to a L1 translation. The Carrollian definition and assessment of memory differs considerably from new conceptions of WM presented in the previous section. A large body of research supports the association between WM and various aspects of L2 learning,¹¹ leading aptitude researchers therefore argue for including WM as a distinct aptitude component (Miyake & Friedman 1998, DeKeyser & Koeth 2011, Robinson 2002, Skehan 2019). Aspects of WM have been integrated into new aptitude conceptions (see §2.1.1 and §2.1.3), such as the Macro SLA-aptitude model (Skehan 2019), the Aptitude Complexes Hypothesis (Robinson 2002) or the Hi-LAB framework (Linck et al. 2013).

There is evidence that WM components relate differentially to language aptitude. A meta-analysis of 66 studies by Li (2016), found EWM to be more strongly associated with aptitude as a whole (moderate correlation of $r = 0.37$) than phonological WM (PWM) with a weak correlation of $r = 0.16$. Li (2016: 828) therefore suggests that EWM is a "more promising aptitude component than PWM". This hypothesis is worth exploring further, especially for young learners who were not included in the meta-analysis by Li (2016).

3.4 Creativity

Language learning and creativity can be associated in two ways. First, current communicative teaching methods, such as the task-based approach (Willis 1996), require learners to contribute their own ideas in order to cope successfully with learning activities. Creative learners may be better equipped to deal with this

¹⁰The meta-analysis by Linck et al. (2014) included data from 79 samples with 3,707 participants and 748 effect sizes.

¹¹For a discussion see e.g. DeKeyser & Koeth (2011), Wen (2015), or Linck et al. (2014) for a meta-analysis on WM and L2 comprehension and production.

kind of learning environment because they are likely to generate ideas easily which leaves them with more mental resources to engage with the target language. Second, creative thinking and language learning are hypothesized to share similar cognitive processing mechanisms, leading to the assumption that creative people are also good language learners, and/or that learning languages is good for creativity (Kharkhurin 2012). For these reasons, creativity has been described as an ID-variable worth exploring in L2 learning and acquisition (Dörnyei & Ryan 2015).

The creativity construct involves a broad range of factors, including cognitive, motivational, personality-linked, societal and procedural aspects which have all been incorporated into different theories of creativity (for an overview see Lubart 1994). In relation to language learning ability, the focus has been narrowed down to the cognitive mechanisms underlying creative thinking. The creative cognition view (Finke et al. 1992, Cropley 2006), which was also adopted in the LAPS project, differs from the conception of creativity as a means of artistic expression. Rather, creativity is seen as a particular way of thinking that is similar to problem-solving skills. It involves two basic thought processes (Guilford 1950): Divergent thinking, i.e. the ability to generate many ideas and convergent thinking, i.e. the ability to pick out a suitable idea and elaborate on it.

The core mechanisms of creative thinking are the ability to successfully retrieve existing knowledge, to focus on important information and suppress the irrelevant, and to analyze and transform this information into novel ideas so that a problem or task can be solved. Individuals must be able to tolerate ambiguity when an answer is not immediately available (Guilford 1950, Finke et al. 1992).

Similar processes are hypothesized to be involved in foreign language learning. The CANAL-F theory (Grigorenko et al. 2000) emphasizes the fact that successful language learners are able to deal well with novelty and tolerate ambiguity in the face of new and unfiltered linguistic input. These individuals can access existing knowledge easily and merge it with new information in order to fill linguistic gaps. People with these abilities are thought to be at the same time creative and good language learners.

Studies exploring creativity and language learning are scarce and have either focused on the possibility that language learning enhances creative thinking, or that creative thinking boosts L2 proficiency. As discussed in Chapter 6, they were mainly conducted with adults or adolescents and have produced mixed findings. Our own work presented in Chapter 6 investigated the effects of creative thinking on L2 proficiency and L2 motivation. To our knowledge, the affective link between creativity and language learning has not previously been considered.

3.5 Cognitive styles – Field independence

Field independence was first mentioned in connection with language aptitude in the 1980s (Chapelle & Green 1992). Originally, this concept was defined as a cognitive style, i.e. a preferred way of cognitively processing information (Witkin et al. 2014). Based on different tasks to assess the perception of verticality, Witkin (1949) identified two conceptualizations of visual processing: Some participants relied on their surroundings as a whole (field dependent) while others perceived individual parts of an image and then reconstructed them (field independent). However, the extent to which a person is field dependent or independent is not a categorical condition but rather located along a continuum. The concept has subsequently been discussed critically by several authors (for an overview, see Evans et al. 2013) and investigated from different angles, namely in connection with WM (Miyake et al. 2001), visual perception (Zhang 2004) or intelligence (Richardson & Turner 2000).

As far as foreign language learning is concerned, different qualities have been associated with field independence: Field dependent learners are thought to benefit from a communicative approach, as they tend to learn from interactions and role models (for an empirical investigation see Johnson et al. 2000). Field independents, on the other hand, may do well when formal aspects of language are focused on, as this caters to their affinity for analytical thinking (see e.g. results from Stansfield & Hansen 1983). Overall, several studies have documented positive effects for field independence on L2 proficiency, even in communicative settings (see for instance Chapelle & Roberts 1986, Carter 1988; or more recently Farsi et al. 2014, Yaghoubi et al. 2014). These tendencies can be illustrated by Chapelle & Green (1992: 59) who suggest that tests for field independence and measures of L2 proficiency usually show “at least a weak, statistically significant, positive correlation” and that field independent individuals “tend to perform better on many types of second-language tests.”

3.6 Metalinguistic awareness

Although metalinguistic awareness was not explicitly part of the initial definition by Carroll, Singleton (2014) emphasizes it as being closely related to language aptitude, especially to grammatical sensitivity and inductive ability (Alderson et al. 1997), which can be subsumed under language-analytic ability (Skehan 1998). While there are various definitions of metalinguistic awareness, briefly stated, it can be described as the ability to “focus on linguistic form and to switch focus between form and meaning” (Jessner 2008: 277). Similarly, language-analytic ability

involves the capacity to reflect on language form as separate from meaning, for example by reasoning analytically about language patterns to arrive at generalizations, as in the PLAB subtest Language Analysis or the MLAT subtest Words in Sentences. Ranta (2002: 163) therefore argues that language-analytic ability and metalinguistic awareness are essentially “two sides of the same coin”. Or as stated by Roehr-Brackin & Tellier (2019: 1111), language-analytic ability “is at the core of the constructs of language learning aptitude and metalinguistic awareness”. In their work with Anglophone children aged 8–9 ($n = 111$), the authors examined language-analytic ability in relation to metalinguistic awareness, suggesting that both significantly predict children’s L2 proficiency, with language analysis being a stronger predictor. The aptitude component of phonetic coding has also been associated with metalinguistic awareness, namely with phonological awareness (Roehr-Brackin & Tellier 2019).

According to Roehr-Brackin & Tellier (2019), the hypothesized link between language aptitude and metalinguistic awareness is substantiated by the observation that different aptitude subcomponents take on different roles in L2 learning as the individual matures. For instance, as discussed in §2.2.3, some findings indicate that younger children draw more strongly on memory while older children rely more on language-analytic ability. This is interpreted by Roehr-Brackin & Tellier (2019) as an indication of developing metalinguistic abilities and literacy skills. However, as has been observed for several issues discussed in this chapter, there is currently not enough empirical evidence to make sound claims, neither about the *evolving memory* versus *language analysis* orientation of children, nor about the relationship between various aptitude components and metalinguistic awareness. The line of work adopted by Roehr-Brackin & Tellier (2019) is therefore worth pursuing in order to clarify the relationship between the two constructs.

4 Affective dispositions: Motivation and related constructs

Affective learner dispositions are among the most thoroughly researched ID variables in SLA and language learning (Ellis 2004: 536). We outline motivation to learn foreign languages (henceforth L2 motivation), and other affective constructs that have been related to L2 achievement, namely L2 anxiety and L2 self-concepts. A final, personality-linked construct we discuss is locus of control.

4.1 L2 motivation

Research into L2 motivation was initiated in the multilingual context of Canada during the late 1950s. Early work investigated how L2 motivation differed from other types of motivation. This resulted in the development of Gardner & Lambert's (1965) socio-educational model of second language acquisition (Gardner 1985: 146, Gardner 2000). It theorizes L2 motivation as being shaped by attitudes toward an L2 speech community and the learner's willingness to integrate into this community. Students are guided by two types of orientations: 1) Integrative orientations, which refer to the desire to learn the language in order to get in contact with and identify with members of the L2 community, and 2) instrumental orientations linked to learning the L2 for some non-linguistic goal (e.g., academic success or social recognition). The former was identified as being more important, and thus, L2 learners with an integrative orientation were expected to be more successful.

Gardner and Lambert's theory triggered extensive research in Canada and beyond (for reviews see e.g. Gardner 1985 or Au 1988), the results of which have been mixed. In the 90s, various scholars challenged Gardner's concept of integrativeness, claiming that the desire to become part of a L2 community is not fundamental for L2 motivation, but applies to specific sociocultural contexts only, such as bilingual cities in Canada, where a specific L2 community is part of the social environment (see e.g., Noels & Clément 1989, Dörnyei 1990, Moïse et al. 1990, Clément et al. 1994).

These critical discussions marked the beginning of a new, more interdisciplinary era which considered theories from other disciplines, such as cognitive and educational psychology. Most notably, Deci & Ryan's (1985, 2002) self-determination theory (SDT) was extended to SLA (see e.g., Dörnyei 1994, Dickinson 1995, Schmidt et al. 1996, Noels et al. 1999, 2000). The central construct in SDT is intrinsic motivation, which subsumes the three basic psychological needs of self-determination, competence, and interpersonal relatedness. An action is intrinsically motivated if it occurs without external pressure and because it is regarded as inherently enjoyable. Extrinsic motivation, in turn, refers to actions that are taken for secondary reasons. These two kinds of motivation are thought to be located on a continuum, where extrinsic forms of motivation can be gradually transformed to intrinsic motivation through the process of internalization (see e.g. Deci & Ryan 1985).

The recognition of SDT as a psychological framework relevant for L2 motivation research was supported by several studies (Noels et al. 2000, Noels 2001). For instance, Noels et al. (2000: 72–74) investigated 159 English-speaking learners of

French. They were able to relate different forms of intrinsic and extrinsic motivation to their counterparts in Gardner's model, integrative and instrumental orientations.

4.2 L2 self-concepts

Global changes that affected mobility and learning contexts led to the abandonment of Gardner's concept of integrativeness. At the same time, social and dynamic aspects of L2 motivation gained in importance. A very influential model that emerged from this trend is Dörnyei's (2005) L2 Motivational Self System (L2MSS), in which traditional constructs are reinterpreted in light of self-theories postulated in the 1980s.¹² In this model, mental future projections of oneself are assumed to trigger motivational forces that guide students in their L2 learning process. Gardner's integrativeness was reconceptualized as the "ideal L2 self", a mental construct which essentially describes the desire to acquire L2 proficiency for personal, social and job-related reasons (Dörnyei 2009).

In parallel to these developments, dynamic system theories (Larsen-Freeman 1997, see also e.g., Ellis & Larsen-Freeman 2006, De Bot et al. 2007, Larsen-Freeman & Cameron 2008, Larsen-Freeman 2017) gained popularity in L2 motivation research. These theories seemed to provide a suitable framework for capturing the complexity, multidimensionality and dynamics of motivational processes in L2 learning (Dörnyei 2010, Waninge 2015). However, empirical research in this area faces serious difficulties in that conventional ways of testing hypotheses using (multiple) regression models with cross-sectional or longitudinal test data are not deemed appropriate for phenomena that are hypothesized to be highly complex and intra-individually dynamic in their time-course (cf. Dörnyei 2014; for methodological considerations see e.g., Verspoor et al. 2011 or Dörnyei et al. 2015).

4.3 L2 anxiety

Foreign language learning anxiety is defined as any negative emotional state in relation to learning and using a foreign language (MacIntyre 1999). It has been closely related to L2 motivation and L2 self-concepts (for a review see Horwitz 2001). Various studies suggest that all of these affective factors mutually influence each other and eventually contribute to success or failure in L2 learning (see e.g., Noels et al. 2000, Pekrun et al. 2002, Stöckli 2004, Kormos & Csizér 2008, Liu &

¹²The L2MSS is particularly based on theories of *possible selves* and *self-discrepancy*. The interested reader is referred to Markus & Nurius (1986) and Higgins (1987), respectively.

Huang 2011, Heinzmann 2013). At the same time, there is no conclusive evidence on the direction of causality, e.g., anxiety might affect learning or be affected by poor learning abilities; in the same way, self-concepts and motivation might be affected by learning ability and learning experiences (Sparks et al. 2011).

4.4 Locus of control

Locus of control has been mentioned as a personality-linked variable relating to L2 learning in the literature (Biedroń 2010, Peek 2016). It describes the extent to which individuals feel in charge of what is happening to them. Locus of control is similar to the concept of self-efficacy described by Bandura (1986) and Rotter (1990) within the social cognitive theory framework. Self-efficacy usually refers to one's self-confidence in particular situations, for instance academic learning, and can therefore change according to context. Locus of control is related to an individual's general tendency to attribute responsibility for outcomes either to internal or external sources. People with internal locus of control tend to believe that they are personally responsible for an outcome. Individuals with external locus of control ascribe their achievements or failures to an external influence. Learners with internal locus of control are expected to attain higher levels of L2 proficiency as they are more likely to take responsibility for their learning.

5 Environmental factors

The influence of environmental factors, such as family and language background or the role of teaching paradigms, are not the main focus of the LAPS project. However, the inquiry into what shapes foreign language learning cannot be done without considering to some extent the interaction between IDs, educational systems and social environment. A sociological view on education provides a complementary view to the psychometric perspective which bears the risk of overemphasizing the individual while neglecting the structures in which the individuals do or do not unfold their potential. This interplay is explored in Chapter 5.

5.1 Family background

Academic development in general and language learning in particular have been shown to be consistently associated with background variables such as parents' educational level, home literacy practices, and the family's socioeconomic features (see e.g., Avineri et al. 2015 for discussion and more references). In particular

the acquisition of (bi-)literacy was and is the object of many studies, and the general pattern in many Western countries shows that educational systems do not consistently even out inequalities in cultural and economic resources present in children's families (see Farkas 2018 for a recent overview and Kigel et al. 2015 for a study in the German-speaking context). Most educational sociologists, inspired by Bourdieu's (1979) influential theory of different types of capital, distinguish at least between two forms of family dispositions: Economic and cultural capital. In the analyses in chapters 4 and 5 we use background variables pertaining to both economic and cultural predispositions of the learners.

5.1.1 Socio-economic family resources

Sociolinguists and sociologists of education have accumulated a great wealth of evidence on the systematic associations of a family's economic wealth and language learning and using. Most of the evidence concerns first or second language learning, studies of the social conditioning of foreign language learning being relatively scarce (but see DESI-Konsortium 2008 for a study that includes social information). A positive association between the socioeconomic status of a child's family and their school performance has been documented extensively (see Entwisle & Alexander 1992 and chapter 5 for more references).

5.1.2 Cultural and educational family resources

Not only parents' economic resources, but also a family's cultural and educational predispositions have been shown to be associated with children's school performance. In the bi- and multilingualism literature, it is generally assumed that parents' own educational background is predictive of the school performance of children in part because of higher or lower affinities of the parents' own experience with education. Therefore, parents' attitudes toward education and their "habitus" is argued to have an important impact on pupils' school performance (Gogolin 1994). Moreover, better educated parents will often also be better prepared to help and support their children in school systems in which learning highly depends on homework tasks.

Moreover, the language repertoires of the families are important resources for additional language learning (Schepens et al. 2016, Schepens et al. 2020), not only in the obvious cases where one of the family languages is the same as a target (foreign) language in school, but also with respect to the general language of instruction and the often cited potential of multilingual children to learn additional languages more easily (as is often assumed to be the case in the multilingualism

literature, e.g. Montanari & Quay 2019; but see Berthele & Udry 2019 for a more critical assessment of the evidence).

Given the prominence of such questions in educational and multilingualism research, it seems important to take into account socioeconomic and cultural factors in a thorough investigation of individual differences in language learning.

5.2 Teaching paradigms

Implementing adequate teaching approaches for young learners has been described as a major challenge in policy making (Garton et al. 2011). Compatible with the perceived global need for communicative skills in English, curricula across the globe have generally come to adopt some form of Communicative Language Teaching (CLT, Krashen 1981, Garton et al. 2011). Implementing these teaching methods can be constrained by local contexts, for instance in terms of resources, cultures of learning or teacher training (Littlewood 2006, Baker 2008). Often, teachers have been found to respond pragmatically with adapting CLT to suit their individual situation (Carless 2003).

Developing communicative language skills is also at the core of the Swiss curriculum. To meet this aim, a task-based approach to language teaching and learning (TBLT) has been adopted (Willis 1996, Ellis 2017). TBLT mediates language through meaningful tasks that are accomplished by using the target language. Learning takes place when students must fill linguistic knowledge gaps encountered during task completion. Language use is therefore elicited by a real communicative need. TBLT can be implemented in different ways, i.e. independent of curricular prescriptions by building solely on learner questions as they arise during task completion, or by drawing on a syllabus that is complementary to the tasks (Ellis 2017).

Swiss teaching manuals are based on TBLT and structured around units on specific topics that are introduced via authentic input. The topics are elaborated on with meaning-focused activities and complemented with elements of explicit vocabulary and grammar teaching. At the end of a unit, learners complete a task that has often a creative focus, such as writing a poem, doing a role play, or painting a picture that is described to the class. In the LAPS project, we were interested in the interplay between L2 proficiency/L2 motivation and this creative element of TBLT (Chapter 6).

6 Summary

In this chapter, we have presented the ID variables and environmental factors that were considered in the LAPS project. They were included in a test battery and questionnaires that were administered to the participants at the beginning of the project. The test results and measures of L2/L3 proficiency provided the basis for addressing several research questions drawn from the literature on IDs and foreign language learning.

Most notably, we explored the underlying structure of the ID variables (Chapter 3) and assessed the predictive value of each variable for L2 proficiency, proposing different models that could be used by teachers to estimate learner potential (Chapter 4). Several issues were addressed in a longitudinal perspective, namely the development of L2/L3 motivation (Chapter 8), common variables underpinning the development of L1 German and L2 English proficiency (Chapter 9) and the dynamics of child language aptitude (Chapter 10). More specific questions concerning environmental factors are addressed in Chapters 5 to 7. We investigated the impact of socioeconomic variables on L2 achievement (Chapter 5), the task-based L2/L3 classroom in relation to creativity (Chapter 6) and the question of whether living close to a French native-speaking community enhanced children's motivation to learn the target language (Chapter 7).

Some of these issues, especially language aptitude, have rarely been studied with children and with large cohorts. As a result, scholarly evidence remains inconclusive and further work is welcome to advance theoretical understanding and methodological innovation in the field in general and with regard to child language aptitude in particular. We hope that our contribution from the LAPS project will add to building a theoretical and pedagogical framework and that it will encourage similar research projects.

References

- Abrahamsson, Niclas & Kenneth Hyltenstam. 2008. The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition* 30(4). 481–509.
- Alderson, J. Charles, Caroline Clapham & David Steel. 1997. Metalinguistic knowledge, language aptitude and language proficiency. *Language Teaching Research* 1(2). 93–121.

- Alexiou, Thomai. 2009. Young learners' cognitive skills and their role in foreign language vocabulary learning. In Marianne Nikolov (ed.), *Second language acquisition: Early learning of modern foreign languages: Processes and outcomes*, 46–61. Bristol: Multilingual Matters.
- Ameringer, Victoria, Luke Green, Daniel Leisser & Sabrina Turker. 2019. Introduction: Towards an interdisciplinary understanding of language aptitude. In Susanne M. Reiterer (ed.), *Exploring language aptitude: Views from psychology, the language sciences, and cognitive science*, vol. 16 (English Language Education), 1–19. Cham: Springer.
- Au, Shun Y. 1988. A critical appraisal of Gardner's social-psychological theory of second-language (L2) learning. *Language Learning* 38(1). 75–99.
- Avineri, Netta, Eric Johnson, Shirley Brice-Heath, Teresa McCarty, Elinor Ochs, Tamar Kremer-Sadlik, Susan Blum, Ana Celia Zentella, Jonathan Rosa, Nelson Flores, H. Samy Alim & Django Paris. 2015. Invited forum: Bridging the “language gap”. *Journal of Linguistic Anthropology* 25. 66–86. DOI: 10.1111/jola.12071.
- Baddeley, Alan D. 2000. The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences* 4(11). 417–423. DOI: 10.1016/S1364-6613(00)01538-2.
- Baddeley, Alan D. 2003. Working memory and language: An overview. *Journal of Communication Disorders* 36(3). 189–208. DOI: 10.1016/S0021-9924(03)00019-4.
- Baddeley, Alan D., Susan Gathercole & Costanza Papagno. 1998. The phonological loop as a language learning device. *Psychological Review* 105(1). 158–173. DOI: 10.1037/0033-295X.105.1.158.
- Baddeley, Alan D. & Graham Hitch. 1974. Working memory. In Gordon H. Bower (ed.), *Psychology of learning and motivation*, 47–89. Cambridge Massachusetts: Academic Press.
- Baker, Will. 2008. A critical examination of ELT in Thailand: The role of cultural awareness. *RELC Journal* 39(1). 131–146. DOI: 10.1177/0033688208091144.
- Bandura, Albert. 1986. *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, N.J.: Prentice-Hall.
- Berthele, Raphael & Isabelle Udry. 2019. Multilingual boost vs. cognitive abilities: Testing two theories of multilingual language learning in a primary school context. *International Journal of Multilingualism*. 1–20. DOI: 10.1080/14790718.2019.1632315.
- Bialystok, Ellen & Maria Fröhlich. 1978. Variables of classroom achievement in second language learning. *The Modern Language Journal* 62(7). 327–336.
- Biedroń, Adriana. 2010. Awareness, responsibility and control as prerequisites of autonomy in second language acquisition. *Beyond Philology* (7). 305–326.

- Birdsong, David. 2014. The Critical Period Hypothesis for second language acquisition: Tailoring the coat of many colors. In Mirosław Pawlak & Larissa Aronin (eds.), *Essential topics in Applied Linguistics and Multilingualism: Studies in honor of David Singleton* (Second Language Learning and Teaching), 43–50. Cham: Springer.
- Bley-Vroman, Robert. 1989. What is the logical problem of foreign language learning? In S. M. Gass & J. Schachter (eds.), *Linguistic perspectives on second language acquisition*, 41–68. New York: Cambridge University Press.
- Bourdieu, Pierre. 1979. *La distinction: Critique sociale du jugement*. Paris: Les Éditions de Minuit.
- Boyle, Whitney, Annukka K. Lindell & Evan Kidd. 2013. Investigating the role of verbal working memory in young children's sentence comprehension. *Language Learning* 63(2). 211–242. DOI: <https://doi.org/10.1111/lang.12003>.
- Carless, David R. 2003. Factors in the implementation of task-based teaching in primary schools. *System* 31. 485–500.
- Carpenter, Helen. 2008. *A behavioral and electrophysiological investigation of different aptitudes for L2 grammar in learners equated for proficiency level*. Washington: Georgetown University. (Doctoral dissertation).
- Carroll, John B. 1958. A factor analysis of two foreign language aptitude batteries. *The Journal of General Psychology* 58(1). 3–19. DOI: 10.1080/00221309.1958.9710168.
- Carroll, John B. 1964. The prediction of success in intensive foreign language training. In *Training research and education*, 87–136. Pittsburgh: University of Pittsburgh Press.
- Carroll, John B. 1973. Implications of aptitude test research and psycholinguistic theory for foreign-language teaching. *Linguistics* 11(112). 5–14.
- Carroll, John B. 1981. Twenty-five years of research on foreign language aptitude. In Karl C. Diller (ed.), *Individual differences and universals in language learning aptitude*, 83–118. Rowley: Newbury House.
- Carroll, John B. & Stanley M. Sapon. 1959. *Modern language aptitude test*. <http://psycnet.apa.org/psycinfo/1960-03629-000>.
- Carroll, John B. & Stanley M. Sapon. 1976. *Modern language aptitude test – Elementary*.
- Carter, Elaine Fuller. 1988. The relationship of field dependent/independent cognitive style to Spanish language achievement and proficiency: A preliminary report. *The Modern Language Journal* 72(1). 21–30.
- Cattell, Raymond B. 1943. The measurement of adult intelligence. *Psychological Bulletin* 40(3). 153–193. DOI: 10.1037/h0059973.

- Chapelle, Carol A. & Pat Green. 1992. Field independence/dependence in second-language acquisition research. *Language Learning* 42(1). 47–83. DOI: 10.1111/j.1467-1770.1992.tb00700.x.
- Chapelle, Carol A. & Cheryl Roberts. 1986. Ambiguity tolerance and field independence as predictors of proficiency in English as a second language. *Language Learning* 36(1). 27–45.
- Christiner, Markus. 2018. Let the music speak: Examining the relationship between music and language aptitude in pre-school children. In Susanne M. Reiterer (ed.), *Exploring language aptitude: Views from psychology, the language sciences, and cognitive neuroscience* (English Language Education), 149–166. Cham: Springer. DOI: 10.1007/978-3-319-91917-1_8.
- Christiner, Markus & Susanne M. Reiterer. 2018. Early influence of musical abilities and working memory on speech imitation abilities: Study with pre-school children. *Brain Sciences* 8(9). 169. DOI: 10.3390/brainsci8090169.
- Clément, Richard, Zoltán Dörnyei & Kimberly A. Noels. 1994. Motivation, self-confidence, and group cohesion in the foreign language classroom. *Language Learning* 44(3). 417–448. DOI: 10.1111/j.1467-1770.1994.tb01113.x.
- Conway, Andrew R. A., Michael J. Kane, Michael F. Bunting, D. Zach Hambrick, Oliver Wilhelm & Randall W. Engle. 2005. Working memory span tasks: A methodological review and user’s guide. *Psychonomic Bulletin & Review* 12(5). 769–786. DOI: 10.3758/BF03196772.
- Cook, Vivian J. 2001. *Second language learning and language teaching*. London: Edward Arnold.
- Cropley, Arthur. 2006. In praise of convergent thinking. *Creativity Research Journal* 18(3). 391–404.
- Daneman, Meredyth & Patricia A. Carpenter. 1980. Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior* 19(4). 450–66.
- De Bot, Kees, Wander Lowie & Marjolijn Verspoor. 2007. A Dynamic Systems Theory approach to second language acquisition. *Bilingualism: Language and Cognition* 10(1). 7–21. DOI: 10.1017/S1366728906002732.
- Deci, Edward L. & Richard M. Ryan. 1985. *Intrinsic motivation and self-determination in human behaviour*. New York: Plenum.
- Deci, Edward L. & Richard M. Ryan (eds.). 2002. *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.
- DeKeyser, Robert. 2000. The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22(4). 499–533.

- DeKeyser, Robert & Joel Koeth. 2011. Cognitive aptitudes for second language learning. In Eli Hinkel (ed.), *Handbook of research in second language teaching and learning*, 395–407. New York: Routledge.
- DESI-Konsortium. 2008. *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie (556610784)*. Tech. rep. Frankfurt am Main: Deutsches Institut für Internationale Pädagogische Forschung (DIPF).
- Dickinson, Leslie. 1995. Autonomy and motivation: A literature review. *System* 23(2). 165–174. DOI: 10.1016/0346-251X(95)00005-5.
- Dörnyei, Zoltán. 1990. Conceptualizing motivation in foreign-language learning. *Language Learning* 40(1). 45–78.
- Dörnyei, Zoltán. 1994. Understanding L2 motivation: On with the challenge! *The Modern Language Journal* 78(4). 515–523.
- Dörnyei, Zoltán. 2005. Motivation and “self-motivation”. In Zoltán Dörnyei (ed.), *The psychology of the language learner individual differences in second language acquisition*, 65–118. Mahwah, NJ: Lawrence Erlbaum Associates.
- Dörnyei, Zoltán. 2009. The L2 motivational self system. In Zoltán Dörnyei & Ema Ushioda (eds.), *Motivation, language identity and the L2 self*. Bristol: Multilingual Matters.
- Dörnyei, Zoltán. 2010. *Questionnaires in second language research*. 2nd edn. New York: Routledge.
- Dörnyei, Zoltán. 2014. Researching complex dynamic systems: ‘Retrodictive qualitative modelling’ in the language classroom. *Language Teaching* 47(1). 80–91. DOI: 10.1017/S0261444811000516.
- Dörnyei, Zoltán, Peter D. MacIntyre & Alastair Henry (eds.). 2015. *Motivational dynamics in language learning*. Bristol, Buffalo, Toronto: Multilingual Matters.
- Dörnyei, Zoltán & Stephen Ryan. 2015. *The psychology of the language learner revisited*. New York: Routledge.
- Doughty, Catherine J. 2013. Optimizing post-critical-period language learning. In Gisela Grañena & Mike Long (eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment*, 153–178. Amsterdam: John Benjamins Publishing.
- Doughty, Catherine J., Susan G. Campbell, Meredith A. Mislevy, Michael F. Bunting, Anita R. Bowles & Joel T. Koeth. 2010. Predicting near-native ability: The factor structure and reliability of Hi-LAB. In Matthew T. Prior, Yukiko Watanabe & Sang-Ki Lee (eds.), *Selected proceedings of the 2008 Second Language Research Forum*, 10–31. Somerville, MA, USA: Cascadia Proceedings Project.
- Ehrman, Madeline E. 1998. The modern language aptitude test for predicting learning success and advising students. *Applied Language Learning* 9(1). 33–70.

- Ellis, Nick C. & Diane Larsen-Freeman. 2006. Language emergence: Implications for applied linguistics. Introduction to the special issue. *Applied Linguistics* 27(4). 558–589. DOI: 10.1093/applin/aml028.
- Ellis, Rod. 2004. Individual differences in second language learning. In Alan Davies & Catherine Elder (eds.), *The handbook of applied linguistics*, 525–551. Oxford: Blackwell Publishing.
- Ellis, Rod. 2017. Task-based language teaching. In Shawn Loewen & Masatoshi Sato (eds.), *The Routledge handbook of instructed second language acquisition*, 108–125. New York: Routledge.
- Entwisle, Doris R. & Karl L. Alexander. 1992. Summer setback: Race, poverty, school composition, and mathematics achievement in the first two years of school. *American Sociological Review* 57(1). 72–84. DOI: 10.2307/2096145.
- Erlam, Rosemary. 2005. Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research* 9(2). 147–171. DOI: 10.1191/1362168805lr1610a.
- Esser, U. & B. Kossling. 1986. A general psychological approach to the diagnosis of foreign language aptitude. In Vivian J. Cook (ed.), *Experimental approaches to second language learning*, 95–100. Oxford: Pergamon.
- Evans, Carol, John T. E. Richardson & Michael Waring. 2013. Field independence: Reviewing the evidence. *British Journal of Educational Psychology* 83(2). 210–224. DOI: 10.1111/bjep.12015.
- Farkas, George. 2018. Family, schooling, and cultural capital. In Barbara Schneider (ed.), *Handbook of the sociology of education in the 21st century*, 3–38. Cham: Springer. DOI: 10.1007/978-3-319-76694-2_1.
- Farsi, Mitra, Mohamed S. Bagheri, Maryam Sharif & Fateme Nematollahi. 2014. Relationship between field dependence/independence and language proficiency of female EFL students. *International Journal of Language Learning and Applied Linguistics World* 6(3). 208–220.
- Finke, Ronald A., Thomas B. Ward & Steven M. Smith. 1992. *Creative cognition*. Massachusetts: MIT Press.
- French, Leif M. & Irena O’Brien. 2008. Phonological memory and children’s second language grammar learning. *Applied Psycholinguistics* 29(3). 463–487. DOI: 10.1017/S0142716408080211.
- Gardner, Howard. 1983. *Frames of mind. The theory of multiple intelligences*. New York: Basic Books.
- Gardner, Robert C. 1985. *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.
- Gardner, Robert C. 2000. Correlation, causation, motivation, and second language acquisition. *Canadian Psychology / Psychologie Canadienne* 41(1). 10–24.

- Gardner, Robert C. & Wallace E. Lambert. 1965. Language aptitude, intelligence, and second-language achievement. *Journal of Educational Psychology* 56(4). 191.
- Garton, Sue, Fiona Copland & Anne Burns. 2011. Investigating global practices in teaching English to young learners. *ELT Research Papers* 11(1). 1–21.
- Gathercole, Susan E. & Susan J. Pickering. 2000. Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *British Journal of Educational Psychology* 70(2). 177–194. DOI: <https://doi.org/10.1348/000709900158047>.
- Gogolin, Ingrid. 1994. *Der monolinguale Habitus der multilingualen Schule*. Münster: Waxmann.
- Golestani, Narly, Cathy J. Price & Sophie K. Scott. 2011. Born with an ear for dialects? Structural plasticity in the expert phonetician brain. *Journal of Neuroscience* 31(11). 4213–4220. DOI: [10.1523/JNEUROSCI.3891-10.2011](https://doi.org/10.1523/JNEUROSCI.3891-10.2011).
- Graaff, Rick de. 1997. The Esperanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition* 19(2). 249–276. DOI: [10.1017/S0272263197002064](https://doi.org/10.1017/S0272263197002064).
- Grañena, Gisela. 2012. *Age differences and cognitive aptitudes for implicit and explicit learning in ultimate second language attainment*. College Park, MA: University of Maryland. (Doctoral Dissertation).
- Grañena, Gisela. 2013. Cognitive aptitudes for second language learning and the LLAMA language aptitude test. In Gisela Grañena & Michael Long (eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (Language Learning and Language Teaching 35), 105–130. Amsterdam: John Benjamins.
- Grañena, Gisela. 2016. Cognitive aptitudes for implicit and explicit learning and information-processing styles: An individual differences study. *Applied Psycholinguistics* 37(3). 577–600.
- Grañena, Gisela & Michael H Long. 2012. Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research* 29(3). 311–343. DOI: [10.1177/0267658312461497](https://doi.org/10.1177/0267658312461497).
- Grigorenko, Elena L., Robert J. Sternberg & Madeline E. Ehrman. 2000. A theory-based approach to the measurement of foreign language learning ability: The Canal-F theory and test. *The Modern Language Journal* 84(3). 390–405.
- Guilford, Joy Paul. 1950. Creativity. *American Psychologist* 5. 444–454.
- Hasselhorn, Marcus & Dietmar Grube. 2003. Das Arbeitsgedächtnis: Funktionsweise, Entwicklung und Bedeutung für kognitive Leistungsstörungen. *Sprache – Stimme – Gehör* (27). 31–37.
- Heinzmann, Sybille. 2013. *Young language learners' motivation and attitudes: Longitudinal, comparative and explanatory perspectives*. London: Bloomsbury.

- Higgins, E. Tory. 1987. Self-discrepancy: A theory relating self and affect. *Psychological Review* 94(3). 319–340.
- Horwitz, Elaine. 2001. Language anxiety and achievement. *Annual Review of Applied Linguistics* 21. 112–126. DOI: 10.1017/S0267190501000071.
- Jessner, Ulrike. 2008. A DST model of multilingualism and the role of metalinguistic awareness. *The Modern Language Journal* 92(2). 270–283. DOI: 10.1111/j.1540-4781.2008.00718.x.
- Johnson, Jacqueline S. & Elissa Newport. 1989. Critical period effects in second language learning: The influence of the maturational state on the acquisition of English as a second language. *Cognitive Psychology* (21). 60–99.
- Johnson, Janice, Suzanne Prior & Mariangela Artuso. 2000. Field dependence as a factor in second language communicative production. *Language Learning* 50(3). 529–567.
- Juffs, Alan & Michael Harrington. 2011. Aspects of working memory in L2 learning. *Language Teaching* 44(2). 137–166. DOI: 10.1017/S0261444810000509.
- Kharkhurin, Anatoliy V. 2012. *Multilingualism and creativity*. Bristol: Multilingual Matters.
- Kigel, Rebecca M., Nele McElvany & Michael Becker. 2015. Effects of immigrant background on text comprehension, vocabulary, and reading motivation: A longitudinal study. *Learning and Instruction* 35. 73–84. DOI: 10.1016/j.learninstruc.2014.10.001.
- Kiss, Csilla. 2009. The role of aptitude in young learners' foreign language learning. In Marianne Nikolov (ed.), *The age factor and early language learning*, 253–276. Berlin: De Gruyter Mouton.
- Kiss, Csilla & Marianne Nikolov. 2005. Developing, piloting, and validating an instrument to measure young learners' aptitude. *Language Learning* 55(1). 99–150.
- Kormos, Judit & Kata Csizér. 2008. Age-related differences in the motivation of learning English as a foreign language: Attitudes, selves, and motivated learning behavior. *Language Learning* 58(2). 327–355. DOI: 10.1111/j.1467-9922.2008.00443.x.
- Kormos, Judit & Anna Sáfár. 2008. Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition* 11(2). 261–271. DOI: 10.1017/S1366728908003416.
- Krashen, Stephen. 1981. *Second language acquisition and second language learning*. Oxford: Pergamon Press. http://sdkrashen.com/content/books/sl_acquisition_and_learning.pdf (6 July, 2018).

- Larsen-Freeman, Diane. 1997. Chaos/complexity science and second language acquisition. *Applied Linguistics* 18(2). 141–165. DOI: 10.1093/applin/18.2.141.
- Larsen-Freeman, Diane. 2017. Complexity theory: The lessons continue. In Lourdes Ortega & ZhaoHong Han (eds.), *Complexity theory and language development. In celebration of Diane Larsen-Freeman* (Language Learning and Language Teaching 48), 11–50. Amsterdam: John Benjamins.
- Larsen-Freeman, Diane & Lynne Cameron. 2008. *Complex systems and applied linguistics*. Oxford: Oxford University Press.
- Lenneberg, Eric Heinz. 1967. *Biological foundations of language*. New York: Wiley.
- Li, Shaofeng. 2016. The construct validity of language aptitude: A meta-analysis. *Studies in Second Language Acquisition* 38(4). 801–842. DOI: 10.1017/S027226311500042X.
- Li, Shaofeng. 2017. The effects of cognitive aptitudes on the process and product of L2 interaction: A synthetic review. In Laura Gurzynski-Weiss (ed.), *Expanding individual difference research in the interaction approach: Investigating learners, instructors, and other interlocutors* (Aila Applied Linguistics Series 16), 41–70. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Li, Shaofeng. 2018. Language aptitude. In Anne Burns & Jack C. Richards (eds.), *The Cambridge guide to learning English as a second language*, 63–72. Cambridge: Cambridge University Press.
- Li, Shaofeng. 2019. Six decades of language aptitude research: A comprehensive and critical review. In Zhisheng (Edward) Wen, Peter Skehan, Adriana Biedroń, Richard L. Sparks & Shaofeng Li (eds.), *Rethinking language aptitude: Contemporary insights and emerging trends*. New York: Routledge.
- Linck, Jared A., Meredith M. Hughes, Susan G. Campbell, Noah H. Silbert, Medha Tare, Scott R. Jackson, Benjamin K. Smith, Michael F. Bunting & Catherine J. Doughty. 2013. Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning* 63(3). 530–566. DOI: 10.1111/lang.12011.
- Linck, Jared A., Peter Osthus, Joel T. Koeth & Michael F. Bunting. 2014. Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review* 21(4). 861–883. DOI: 10.3758/s13423-013-0565-2.
- Littlewood, William. 2006. Communicative and task-based language teaching in East Asian classrooms. *Language Teaching* 40. 243–349.
- Liu, Meihua & Wenhong Huang. 2011. An exploration of foreign language anxiety and English learning motivation. *Education Research International* 2011. 1–8.
- Lubart, Todd I. 1994. Creativity. In Robert Sternberg (ed.), *Handbook of perception and cognition. Thinking and problem solving*, 290–323. Cambridge, Massachusetts: Academic Press.

- MacIntyre, Peter D. 1999. Language anxiety: A review of the research for language teachers. In Dolly Jesusita Young (ed.), *Affect in foreign language and second language teaching: A practical guide to creating a low-anxiety classroom atmosphere*, 24–45. Boston, MA: McGraw-Hill.
- Markus, Hazel & Paula Nurius. 1986. Possible selves. *American Psychologist* 41(9). 954–969.
- Meara, Paul M., James Milton & Nuria Lorenzo-Dus. 2001. *Llama language aptitude tests: The manual*. Swansea.
- Milton, James & Thomai Alexiou. 2006. Language aptitude development in young learners. In C. Abello-Contesse, R. Chacón-Beltrán & M. D. López-Chiménez (eds.), *Age in L2: Acquisition and teaching*, 177–192. New York: Peter Lang.
- Miyake, Akira & Naomi P. Friedman. 1998. Individual differences in second language proficiency: Working memory as language aptitude. In *Foreign language learning: Psycholinguistic studies on training and retention*, 339–364. Mahwah, NJ: Lawrence Erlbaum Associates.
- Miyake, Akira, Alexander H. Witzki & Michael J. Emerson. 2001. Field dependence–independence from a working memory perspective: A dual-task investigation of the Hidden Figures Test. *Memory* 9(4-6). 445–457. DOI: 10.1080/09658210143000029.
- Moïse, Léna Céline, Richard Clément & Kimberly Noels. 1990. Aspects motivationnels de l'apprentissage de l'espagnol au niveau universitaire. *Canadian Modern Language Review* 46(4). 689–705. DOI: 10.3138/cmlr.46.4.689.
- Montanari, Simona & Suzanne Quay. 2019. Facilitated language learning in multilinguals. In Simona Montanari & Suzanne Quay (eds.), *Multidisciplinary perspectives on multilingualism: The fundamentals*, 299–324. Berlin: De Gruyter Mouton. DOI: 10.1515/9781501507984.
- Morgan-Short, Kara, Mandy Faretta-Stutenberg, Katherine A. Brill-Schuetz, Helen Carpenter & Patrick C. M. Wong. 2014. Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition* 17(1). 56–72. DOI: 10.1017/S1366728912000715.
- Muñoz, Carmen. 2014. The association between aptitude components and language skills in young learners. In Mirosław Pawlak & Larissa Aronin (eds.), *Essential topics in applied linguistics and multilingualism*, 51–68. Cham: Springer.
- Noels, Kimberly A. 2001. New orientations in language learning motivation: towards a model of intrinsic, extrinsic, and integrative orientations and motivations. In Zoltán Dörnyei & Richard Schmidt (eds.), *Motivation and second language acquisition*, 43–68. Honolulu: University of Hawaii Press.

- Noels, Kimberly A. & Richard Clément. 1989. Orientations to learning German: The effects of language heritage on second-language acquisition. *Canadian Modern Language Review* 45(2). 245–257. DOI: 10.3138/cmlr.45.2.245.
- Noels, Kimberly A., Richard Clément & Luc G. Pelletier. 1999. Perceptions of teachers' communicative style and students' intrinsic and extrinsic motivation. *The Modern Language Journal* 83(1). 23–34. DOI: 10.1111/0026-7902.00003.
- Noels, Kimberly A., Luc G. Pelletier, Richard Clément & Robert J. Vallerand. 2000. Why are you learning a second language? Motivational orientations and self-determination theory. *Language Learning* 50(1). 57–85.
- O'Brien, Irena, Norman Segalowitz & Joe Collentine. 2006. Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics* (27). 377–402.
- Ottó, István. 1996. Language aptitude testing: Unveiling the mystery. *Novelty* 3(3). 6–20.
- Parry, Thomas S. & James R. Child. 1990. Preliminary investigation of the relationship between VORD, MLAT and language proficiency. In Thomas S. Parry & Charles W. Stansfield (eds.), *Language aptitude reconsidered*, 30–67. New Jersey: Center for Applied Linguistics.
- Pashler, Harold, Mark McDaniel, Doug Rohrer & Robert Bjork. 2008. Learning styles: Concepts and evidence. *Psychological Science in the Public Interest* 9(3). 105–119. DOI: 10.1111/j.1539-6053.2009.01038.x.
- Peek, Ron. 2016. Exploring learner autonomy: language learning locus of control in multilinguals. *International Journal of Multilingualism* 13(2). 230–248. DOI: 10.1080/14790718.2015.1090991.
- Pekrun, Reinhard, Thomas Goetz, Wolfram Titz & Raymond P. Perry. 2002. Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist* 37(2). 91–106.
- Petersen, Calvin R. & Antoine R. Al-Haik. 1976. The development of the Defense Language Aptitude Battery (Dlab). *Educational and Psychological Measurement* 36(2). 369–380. DOI: 10.1177/001316447603600216.
- Pimsleur, Paul. 1966. *Pimsleur language aptitude battery (form S.)* San Diego: Harcourt Brace, & Co.
- Pimsleur, Paul & Terence Quinn. 1971. *The psychology of second language learning: Papers from the second International Congress of Applied Linguistics, Cambridge, 8-12 September*. Terence Quinn (ed.). <http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=ED058773>.
- Plomin, Robert. 2019. *Blueprint, with a new afterword: How DNA makes us who we are*. London: Penguin Books.

- Ranta, Leila. 2002. The role of learners' language analytic ability in the communicative classroom. In Peter Robinson (ed.), *Individual differences and instructed language learning*, 159–180. Amsterdam: Benjamins.
- Ranta, Leila. 2008. Aptitude and good language learners. In Carol Griffiths (ed.), *Lessons from good language learners*, 142–156. New York: Cambridge University Press.
- Reiterer, Susanne M., Xiaochen Hu, T. A. Sumathi & Nandini Chatterjee Singh. 2013. Are you a good mimic? Neuro-acoustic signatures for speech imitation ability. *Frontiers in Psychology* 4. DOI: 10.3389/fpsyg.2013.00782.
- Richardson, Julie A. & Tony E. Turner. 2000. Field dependence revisited I: Intelligence. *Educational Psychology* 20(3). 255–270. DOI: 10.1080/713663747.
- Riener, Cedar & Daniel Willingham. 2010. The myth of learning styles. *Change: The Magazine of Higher Learning* 42(5). 32–35. DOI: <https://doi.org/10.1080/00091383.2010.503139>.
- Rimfeld, Kaili, Philip S. Dale & Robert Plomin. 2015. How specific is second language-learning ability? A twin study exploring the contributions of first language achievement and intelligence to second language achievement. *Translational Psychiatry* 5(9). 638. DOI: 10.1177/0956797610368060.
- Robinson, Peter. 1997. Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning* 47(1). 45–99.
- Robinson, Peter. 2001. Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics* 22(1). 27–57. DOI: 10.1093/applin/22.1.27.
- Robinson, Peter. 2002. Learning conditions, aptitude complexes and SLA: A framework for research and pedagogy. In *Individual differences and instructed language learning*, 113–136. Amsterdam: Benjamins.
- Roehr-Brackin, Karen & Angela Tellier. 2019. The role of language-analytic ability in children's instructed second language learning. *Studies in Second Language Acquisition* 41. 1111–1131.
- Rogers, Vivienne, Paul Meara, Thomas Bennett-Legh, Clare Curry & Emma Davie. 2017. Examining the LLAMA aptitude tests. *Journal of the European Second Language Association* 1(1). 49–60. <https://www.euroslajournal.org/articles/10.22599/jesla.24/> (12 January, 2021).
- Rotter, Julian B. 1990. Internal versus external control of reinforcement: A case history of a variable. *American Psychologist* 45(4). 489–493.
- Sáfár, Anna & Judit Kormos. 2008. Revisiting problems with foreign language aptitude. *International Review of Applied Linguistics in Language Teaching* 46(2). 113–136.

- Sasaki, Miyuki. 1996. *Second language proficiency, foreign language aptitude, and intelligence: Quantitative and qualitative analyses*. New York: Peter Lang.
- Sattler, Jerome M. (ed.). 2001. *Assessment of children's cognitive applications*. La Mesa, California: Sattler Publisher Inc.
- Schepens, Job, Frans van der Slik & Roeland van Hout. 2016. L1 and L2 distance effects in learning L3 Dutch. *Language Learning* 66(1). 224–256. DOI: <https://doi.org/10.1111/lang.12150>.
- Schepens, Job, Roeland van Hout & T. Florian Jaeger. 2020. Big data suggest strong constraints of linguistic similarity on adult language learning. *Cognition* 194. 104056. DOI: [10.1016/j.cognition.2019.104056](https://doi.org/10.1016/j.cognition.2019.104056).
- Schmidt, Richard, Deena Boraie & Omneya Kassabgy. 1996. *Foreign language motivation: Internal structure and external connections*. Working Paper. <http://scholarspace.manoa.hawaii.edu/handle/10125/40769>.
- Singleton, David. 2014. Apt to change: The problematic of language awareness and language aptitude in age-related research. *Studies in Second Language Learning and Teaching* 4(3). 557–571.
- Singleton, David. 2017. Language aptitude: Desirable trait or acquirable attribute? *Studies in Second Language Learning and Teaching* 7(1). 89–103.
- Skehan, Peter. 1986. Cluster analysis and the identification of learner types. In Vivian J. Cook (ed.), *Experimental approaches to second language acquisition*, 81–94. Oxford: Pergamon Press.
- Skehan, Peter. 1998. *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, Peter. 2002. Theorising and updating aptitude. In Peter Robinson (ed.), *Individual differences and instructed language learning*, 69–93. Amsterdam/Philadelphia: John Benjamins.
- Skehan, Peter. 2019. Language aptitude implicates language and cognitive skills. In Peter Skehan, Adriana Biedroń, Richard L. Sparks, Zhisheng (Edward) Wen & Shaofeng Li (eds.), *Language aptitude advancing theory, testing, research and practice*, 53–75. New York: Routledge.
- Skehan, Peter & Lucile Ducroquet. 1988. *A comparison of first and foreign language learning ability*. <http://www.opengrey.eu/item/display/10068/509355>.
- Snow, Richard. 1991. Aptitude-Treatment Interaction as a framework for research on individual differences in psychotherapy. *Journal of Consulting and Clinical Psychology* 59(2). 205–216.
- Sparks, Richard L. & Leonore Ganschow. 1991. Foreign language learning differences: Affective or native language aptitude differences? *The Modern Language Journal* 75(1). 3–16. DOI: [10.2307/329830](https://doi.org/10.2307/329830).

- Sparks, Richard L., John Patton, Leonore Ganschow & Nancy Humbach. 2011. Sub-components of second-language aptitude and second-language proficiency. *The Modern Language Journal* 95. 253–273.
- Spearman, Charles. 1904. General intelligence, objectively determined and measured. *American Journal of Psychology* 15. 201–293.
- Speciale, Giovanna, Nick C. Ellis & Tracey Bywater. 2004. Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics* (25). 293–321.
- Spolsky, Bernard. 1995. Prognostication and language aptitude testing. *Language Testing* 12(3). 321–340.
- Stansfield, Charles W. 1989. Language aptitude reconsidered. *Eric Digest*. <https://eric.ed.gov/?id=ED318226> (24 March, 2020).
- Stansfield, Charles W. & Jacqueline Hansen. 1983. Field dependence-independence as a variable in second language cloze test performance. *TESOL Quarterly* 17(1). 29–38.
- Stansfield, Charles W. & Daniel J. Reed. 2004. The story behind the Modern Language Aptitude Test: An interview with John B. Carroll (1916–2003). *Language Assessment Quarterly: An International Journal* 1(1). 43–56.
- Stansfield, Charles W. & Daniel J. Reed. 2005. *Modern Language Aptitude Test – Elementary: Spanish Version – MLAT-ES*.
- Stansfield, Charles W. & Daniel J. Reed. 2019. The MLAT at 60 Years. In Zhisheng (Edward) Wen, Peter Skehan, Adriana Biedroń, Richard L. Sparks & Shaofeng Li (eds.), *Rethinking language aptitude: Contemporary insights and emerging trends*. New York: Routledge.
- Sternberg, Robert J. 1985. *Beyond IQ: A triarchic theory of human intelligence*. Cambridge: CUP Archive.
- Sternberg, Robert J. 2002. The theory of successful intelligence and its implications for language aptitude testing. In Peter Robinson (ed.), *Individual differences and instructed language learning*, vol. 2, 13–44. Amsterdam: Benjamins.
- Stöckli, Georg. 2004. *Motivation im Fremdsprachenunterricht: Eine theoriegeleitete empirische Untersuchung in 5. und 6. Primarschulklassen mit Unterricht in Englisch und Französisch* (Pädagogik bei Sauerländer). Oberentfelden: Sauerländer Verlag.
- Stromswold, Karin. 2001. The heritability of language: A review and meta-analysis of twin, adoption, and linkage studies. *Language* 77(4). 647–723. DOI: 10.1353/lan.2001.0247.
- Suárez Vilagran, Maria Del Mar. 2010. *Language aptitude in young learners: The elementary modern language aptitude test in Spanish and Catalan*. Barcelona: Universitat de Barcelona. (Doctoral Dissertation).

- Tellier, Angela & Karen Roehr-Brackin. 2017. Raising children's metalinguistic awareness to enhance classroom second language learning. In María del Pilar García Mayo (ed.), *Learning foreign languages in primary school*, 22–48. Bristol, Blue Ridge Summit: Multilingual Matters. DOI: 10.21832/9781783098118-004.
- Tomasello, Michael. 2005. *Constructing a language: A usage-based theory of language acquisition*. Harvard: Harvard University Press.
- Turner, Marilyn L & Randall W Engle. 1989. Is working memory capacity task dependent? *Journal of Memory and Language* 28(2). 127–154. DOI: 10.1016/0749-596X(89)90040-5.
- Vanhove, Jan. 2013. The critical period hypothesis in second language acquisition: A statistical critique and a reanalysis. *PLOS One* 8(7). DOI: <https://doi.org/10.1371/journal.pone.0069172>.
- Verspoor, Marjolijn H., Kees de Bot & Wander Lowie (eds.). 2011. *A dynamic approach to second language development: Methods and techniques*. Amsterdam: Benjamins.
- Waninge, Frea. 2015. Motivation, emotion and cognition: Attractor states in the classroom. In Zoltán Dörnyei (ed.), *Motivational dynamics in language learning* (Second language acquisition 81), 195–213. Bristol: Multilingual Matters.
- Wen, Zhisheng (Edward). 2015. Working memory in second language acquisition and processing: The phonological/executive model. In Zhisheng (Edward) Wen, Mailce Borges Mota & Arthur McNeill (eds.), *Working memory in second language acquisition and processing*, 41–62. Bristol: Multilingual Matters.
- Wen, Zhisheng (Edward). 2019. Working memory as language aptitude: The Phonological/Executive Model. In Zhisheng (Edward) Wen, Peter Skehan, Adriana Biedroń, Richard L. Sparks & Shaofeng Li (eds.), *Language aptitude: Advancing theory, testing, research and practice*. New York: Routledge.
- Wen, Zhisheng (Edward), Adriana Biedroń & Peter Skehan. 2017. Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching* 50(1). 1–31. DOI: 10.1017/S0261444816000276.
- Wen, Zhisheng (Edward), Peter Skehan, Adriana Biedroń, Shaofeng Li & Richard L. Sparks (eds.). 2019. *Language aptitude: Advancing theory, testing, research and practice* (Second Language Acquisition and Research Series). New York: Routledge.
- Wesche, Marjorie. 1981. Language aptitude measures in streaming, matching students with methods, and diagnosis of learning problems. In Karl C. Diller (ed.), *Individual differences and universals in language learning aptitude*, 119–154. Rowley: Newbury House Publishers.

- Wesche, Marjorie, Henry Edwards & Winston Wells. 1982. Foreign language aptitude and intelligence. *Applied Psycholinguistics* 3(2). 127–140. DOI: 10.1017/S0142716400006664.
- Williams, J. N. 1999. Memory, attention, and inductive learning. *Studies in Second Language Acquisition* 21(01). 1–48.
- Willis, Jane. 1996. A flexible framework for task-based learning. In David Willis & Jane Willis (eds.), *Challenge and change in language teaching*, 52–62. New York: Macmillan Education.
- Winke, P. 2013. An investigation into second language aptitude for advanced Chinese language learning. *The Modern Language Journal* 97(1). 109–130.
- Witkin, Herman A. 1949. Perception of body position and of the position of the visual field. In *Psychological Monographs: General and Applied*, vol. 63, 1–46. Massachusetts: The American Psychological Association.
- Witkin, Herman A., Philip K. Oltman, Evelyn Raskin, Stephen A. Karp & Jack Demick. 2014. *Group embedded figures test manual*. Menlo Park: Mind Garden.
- Yaghoubi, Maryam, Azam Inanloo & AliAkbar Hamed. 2014. Field dependence-independence, cognitive styles, English reading comprehension skill, English language learning, and educational achievement. *Journal of Educational and Management Studies* 4(2). 460–464.
- Zhang, Li-fang. 2004. Field-dependence/independence: Cognitive style or perceptual ability? Validating against thinking styles and academic achievement. *Personality and Individual Differences* 37(6). 1295–1311. DOI: 10.1016/j.paid.2003.12.015.

Chapter 2

Language Aptitude at Primary School (LAPS): Research design

Carina Steiner^a, Raphael Berthele^b & Isabelle Udry^{b,c}

^aUniversity of Bern, Center for the Study of Language and Society ^bUniversity of Fribourg, Institut de Plurilinguisme ^cZurich University of Teacher Education

This chapter delineates the design of the LAPS project. We start with an outline of the research questions, followed by a description of the curricular context of foreign language learning at Swiss primary schools. Next, we will give a full description of the test instruments and how they were implemented, as well as details on the participants and procedures. Finally, data entry and scoring will be outlined.

1 Research questions

The aim of the project *Language Aptitude at Primary School* (LAPS) is to explore the extent to which skills, abilities, and socio-environmental factors contribute to successful language learning by primary school children.

We consider individual difference (ID) variables and environmental factors previously found to affect foreign language learning along four broad categories:

1. Language aptitude according to Carroll (1958): Phonetic coding ability, grammatical sensitivity, inductive ability, rote memory;
2. General cognitive abilities or general learning abilities: Intelligence, working memory, creativity, field independence;
3. Affective dispositions: L2/L3 motivation, L2/L3 self-concepts, foreign language learning anxiety, dedication, perceived teacher encouragement, perceived parental encouragement, locus of control;



4. Environmental factors: Socio-economic status (SES), language background, region, teaching paradigm (i.e. task-based teaching and learning as prescribed by the Swiss curriculum).

The theoretical underpinnings of these variables are discussed in Chapter 1. In the LAPS project, we address the following research questions:

1. What are the underlying dimensions of the ID variables assessed in the test battery and how do these dimensions relate to L2 proficiency? (Chapter 3)
2. What is the predictive value of the ID variables and environmental factors for young learners' L2 proficiency? (Chapter 4)
3. What is the relationship between the ID variables, environmental factors, and foreign language skills?
 - Association of socio-environmental background variables with foreign language ability (Chapter 5)
 - Creativity and the task-based learning environment (Chapter 6)
 - Affective predispositions and the proximity to the target language community (Chapter 7)
4. What are the developmental patterns identified over two academic years (1.5 years of foreign language learning) in terms of:
 - The dynamics of affective dispositions (Chapter 8)
 - Competences in school language German and L2 English (Chapter 9)
 - The language analysis component of language aptitude, i.e. grammatical sensitivity and inductive ability (Chapter 10)

2 Study context

The research questions were investigated in two subprojects between 2017 and 2019. LAPS I took place in a German-speaking region close to the French-German language border. The children learnt L2 French and L3 English. It was designed as a cross-sectional study which was complemented by a second data collection to further investigate affective dispositions and L3 proficiency. LAPS I also served as a pilot study for the test battery.

LAPS II was conducted in the north-eastern German-speaking part of Switzerland. Participants learnt L2 English and L3 French. It was longitudinal with three

data collections (T1, T2, T3) over two academic years. We followed the development of L2 English proficiency, L1 school language proficiency in German, aptitude (language analysis component), and affective dispositions.

2.1 The foreign language curriculum in Switzerland

All Swiss children learn two foreign languages as part of the mandatory curriculum: English and one national language, i.e. French, Italian, or Romansh. The cantons (equivalent to provinces or districts) are free to organise their own language curricula, although, based on a constitutional article accepted by the people in 2006, the last decades were characterized by a tendency to harmonize curricula across the country. Despite this tendency, cantons are still free to choose which two languages they want children to learn and in what order. Overall, this has resulted in two different systems across the German-speaking part of Switzerland where this study took place:

- a. Regions close to the French-speaking area introduce French in 3rd grade and English in 5th grade. Similarly, regions close to the Italian and Romansh speaking parts of the country choose Italian and/or Romansh as one of their languages.
- b. The other regions start with English as the first foreign language in 3rd grade and introduce a national language later, e.g. French, in 5th grade.¹

As mentioned before, in the current project, both systems are represented. LAPS I took place in region a) with L2 French and L3 English, and LAPS II in region b) with L2 English and L3 French.

Each foreign language is taught for 2 to 3 lessons a week, depending on the grade. Tables 5 and 8 indicate the total number of lessons children in the LAPS project had attended at different times of testing.

2.2 Overall goals of the Swiss foreign language curriculum

Foreign language teaching in Switzerland aims at developing functional multilingualism. All four skills are taught in the L2 and L3 from the start: Listening, reading, speaking, and writing. The curriculum also contains a domain called *Sprachen im Fokus* (languages in focus) which covers formal aspects of language

¹The region where LAPS II was conducted is an exception: Until 2019 L2 English was introduced in 2nd grade. In 2020 the region adapted to the common national practice, i.e. L2 English classes now start in 3rd grade.

(including grammar and pronunciation), language awareness, and the use of strategies. *Kulturen im Fokus* (culture in focus) stipulates objectives of cultural knowledge and attitudes. In keeping with communicative approaches to language teaching and learning, fluency is given priority over accuracy. At the end of primary school (age 12), children should reach beginner levels in the L2 and L3. More specifically, the national standards based on the CEFRL² target L2 levels of A2.1 in listening, reading and speaking, and A1.2 in writing. In the L3, children are expected to reach A1.2 levels in all areas of competence (Bildungsdirektion Kanton Zürich 2017: 17).

Children are taught in a task-supported approach, i.e. a weak form of communicative task-based learning and teaching (Ellis 2017). This means that tasks are central to the lesson, but they are usually complemented with form-focused elements. As described in more detail in Chapter 6, this method is reflected in the teaching manuals which in most cantons are prescribed by the local board of education. The teaching manuals contain several units structured around a topic that is introduced via authentic input followed by meaning-focused activities. Vocabulary learning and some elements of explicit grammar are also part of the lesson plans. At the end of each unit, learners use all aspects of language they have acquired to complete a communicative task about the topic, such as writing a poem, or doing role plays.

The Swiss curriculum prescribes awareness raising elements that draw on all languages in a student's repertoire (EDK 2004, Passepartout, Arbeitsgruppe Rahmenbedingungen 2008). The aim is to enhance learning under relatively limited input conditions by developing metalinguistic awareness. Therefore, teaching manuals include specific sections on intercomprehension and language learning strategies that should encourage transfer among languages. However, these sections are not the main focus of the manuals and teachers therefore integrate them flexibly into their lessons (Bildungsdirektion Kanton Zürich 2017: 8).

3 Test battery

The ID variables were assessed in a comprehensive test battery, including psychometric tests and two student questionnaires (motivation, locus of control) and a parent questionnaire (family background information). Where possible, the constructs were measured with standardized tools. However, some tests had to be translated into German and/or adapted to the age of our participants. In the following, short descriptions of each test are given. Tables 1–4 summarize dimen-

²Common European Framework of Reference for Languages

sions, conditions for administration, and in which subproject the tests were used. The reliability analysis is available in the technical report at <https://osf.io/hstv7/>.

The test battery was trialled in LAPS I with 10 classes of 4th and 5th graders from 9 different schools. Some changes were added in consultation with the scientific advisory board of the project before LAPS II (see 3.2).

3.1 Tests for language aptitude

Team up Words! Test of grammatical sensitivity based on the MLAT-E, part 2 Matching Words (Carroll & Sapon 2010), translated into German and adapted for the target group of the present study.

Participants are instructed to identify functions of words in sentences (no explicit grammatical terms are used). After the training phase, they are presented with paired sentences. In the first sentence, one word is highlighted. The participants' task is to find the corresponding word (i.e. the word with the same grammatical function) in the second sentence.

Language Detective Test for inductive abilities based on PLAB form 4 (Pimsleur et al. 2004), translated and adapted for the target group of the present study.

Participants are presented with a list of words and short sentences in an artificial language as well as their translation in German (=language of instruction). From this input, participants have to deduce how sentences in the artificial language may be formed.

Llama-E Test for sound-symbol association and phonemic working memory (Meara et al. 2001). In the training phase, participants learn how sounds are represented graphically in an artificial language. In the test phase, they listen to bi-syllabic words in the artificial language and have to choose the correct "spelling" between two options.

Llama-D Sound recognition task (Meara et al. 2001). In the training phase, participants hear different sounds of an artificial language. In the test phase, participants listen to strings of sounds (combined set of training and novel sounds) and are asked to identify the sounds they have already heard in the training.

3.2 Tests for cognition/general learning abilities

CFT 20-R: Matrices and topological deductions (Weiß 2006): Fluid intelligence was assessed with two subtests from the CFT 20-R (Weiß 2006). In the matrices subtests, geometric patterns with a missing piece are presented to the participants. They are given five choices to pick from and fill in the missing piece. The topologies deductions subtest consists of geometric pattern containing one to three dots. Participants are asked to select, from five given options, the one that replicates the conditions in the example (e.g. the one where a dot could be placed outside a triangle but inside a circle).

CFT-20-R: Number sequences (Weiß 2006): In this test for crystallized intelligence, participants are presented with a sequence of numbers in a certain pattern and are asked to choose the next logical number from five options.

Corsi Blocks: The Corsi Blocks task assesses visual working memory. An increasingly long sequence of squares on the screen is lit up, and participants are asked to reproduce the correct order.

Digit Span (Forward/Backward): Number strings of increasing length are presented visually and aurally. The participants have to reproduce them either in the order of presentation (forward) or in reversed order (backward) until they reach maximum recall capacity. The Digit Span task measures verbal working memory. The forward version is a simple measure of short-term phonological memory (or phonological loop), and the Backward Digit Span is a complex task tapping into executive working memory.

Alphabet Task: A measure of automatic letter access, retrieval, and production by Berninger et al. (1992) in which participants are instructed to write down the alphabet as quickly as possible without sacrificing legibility. The Alphabet task has been associated with children's levels of composition in the L1 (Berninger et al. 1997, Graham et al. 2006).

Test of Creative Thinking (Divergent Production) (TCT-DP) (Urban & Jellen 1995): Participants are presented with a sheet of paper containing a frame with 6 figural fragments (a semicircle, waveline, dot, right angle, dashed line, and a lying "u" located outside the frame). They are asked to complete the unfinished picture without a time limit. The test is scored according to 14 criteria which include boundary breaking, risk taking, introducing new elements, humour, and the ability to link elements to give meaning to the overall picture. Thus, the test goes beyond scoring quantitative

aspects of divergent thinking, such as the number of ideas. The test taps more holistically into an individual's creative potential by considering qualitative aspects of divergent thinking as well, i.e. integrating various elements meaningfully or unexpected ideas.

Group Embedded Figures Test (GEFT): An assessment of field independence (Witkin et al. 2014) in which participants need to find simple geometrical figures embedded in more complex figures.

3.3 Assessment of affective dispositions

Based on existing test instruments from Horwitz et al. (1986), Stöckli (2004), Dörnyei (2010), Heinzmann (2013), and Peyer et al. (2016), we put together a student questionnaire covering the following dimensions: Intrinsic motivation, extrinsic motivation (school/leisure), lingua franca motivation, foreign language learning anxiety, self-concepts (L2 + school language), teacher motivation, parental encouragement, dedication, and future L2 self. The questionnaire comprised a section for L2 and L3 with the same items for each language.

Locus of control was assessed with a German translation of the N-S Personality Scale by Nowicki & Strickland (1973).

3.4 Assessment of environmental factors

A parent questionnaire filled in at the beginning of the study assessed personal and linguistic background (country of origin, years of schooling, L1, family language, literacy language), SES (parents' highest level of education, n° of books, financial resources, monthly income), and school context (classes in German as a second language/heritage language and culture, French/English homework).

3.5 Tests for language proficiency

ELFE 1–6, Reading Proficiency (Lenhard & Schneider 2006) is a normed test for reading skills in the language of instruction German. Items are presented at word, sentence, and text level.

Oxford Young Learners Placement Test (Oxford English Testing 2013) was used for L2/L3 English. The test consists of two sections: Language use (vocabulary and grammar) and listening (short and extended listening exercises) and is said by the distributors to cover levels A1–B1.

C-tests were used for L2 French and L3 English. Participants need to reconstruct meaning from partly deleted words in a short text, completing the missing part of the words. C-tests measure general language proficiency which is

conceptualized as an underlying ability consisting of knowledge and skills displayed in all areas of language use (Eckes & Grotjahn 2006). The C-tests were based on topics and vocabulary covered in the curriculum. They were piloted with classes who did not participate in the LAPS project. For L2/L3 English, texts were adapted from Babaii & Shahri (2010) and Porsch & Wilden (2017) in accordance with curricular content of the target group.

3.6 Piloting and adapting the test battery

After piloting the test battery in LAPS I with 10 classes of 4th and 5th graders from nine different schools, the following changes were added in consultation with the scientific advisory board. The adapted version was used in LAPS II.

3.6.1 L2 English proficiency measure

For LAPS II, the dependent variable for L2 proficiency needed to be changed from L2 French to L2 English, due to curricular differences in the regions of LAPS I and LAPS II outlined in §2.1. This change had been anticipated and tests for L2 English had been selected at the start of the project.

For L2 English proficiency, we initially planned to use the Oxford Young Learners Placement Test (OYLPT, Oxford English Testing 2013). This is an online test assessing L2 English listening comprehension and language use (vocabulary and grammar) embedded in communicative situations. The OYLPT is supposed to cover CEFR levels A0 to B1. The test seemed appropriate for two reasons: First, items are focused on communicative aspects of language use which is in keeping with curricular goals set for the target group. Second, because our participants are expected to reach A2 levels by the end of primary school, we assumed that the OYLPT was suitable to cover the range of proficiency levels in our sample.

We were therefore surprised to find a large group of participants reaching close to the maximum score at the first time of testing (T1). To avoid ceiling effects at T2 and T3, we decided to change the English proficiency measure. In hindsight, it would have been important to pilot the OYLPT with a sample of learners comparable to our LAPS-II-learners as part of the trial phase.

C-tests were chosen as an alternative because they have been shown to be a time-efficient and reliable measure of general language proficiency (Eckes & Grotjahn 2006). Since the Swiss curriculum fosters all communicative skills, including reading and writing, participants have acquired the skills needed to cope with C-tests.

2 Language Aptitude at Primary School (LAPS): Research design

Table 1: Description of language aptitude tests

Subdimension	Test	Conditions	Study
Grammatical sensitivity	Team up Words! (adapted from MLAT-E subtest Matching Words, Carroll & Sapon 2010)	30 items, 13min 30s	LAPS I&II
Inductive ability	Language Detective (adapted from PLAB 4, Pimsleur et al. 2004)	15 items, 10min	LAPS I&II
Sound-symbol association (phonemic working memory)	Llama-E (Meara et al. 2001)	Training: 2min., 24 sounds to learn, Test: 20 bi-syllabic words	LAPS I
Sound recognition task (phonemic discrimination)	Llama-D (Meara et al. 2001)	Training: 10 words, Test: learned words alongside 20 novel words	LAPS II

Table 2: Description of tests for affective dispositions and environmental factors

Subdimension	Test	Conditions	Study
Affective dispositions			
Motivation Foreign language learning anxiety L2/L3 self-concepts	Student questionnaire	LAPS I: T1: 40 items (L2 French), T2: 72 items (L2 French + L3 English), LAPS II: T1: 40 items (L2 English), T2-3: 72 items (L2 English + L3 French) 4-point likert scale, no time constraints	LAPS I&II
Locus of control	N-S personality scale (Nowicki & Strickland 1973)	20 yes/no items	LAPS I&II
Environmental factors			
SES Language background	Parental questionnaire	20 items	LAPS I&II

Table 3: Description of tests for cognition/general learning abilities

Subdimension	Test	Conditions	Study
Fluid intelligence	CFT 20-R: Matrices (Weiß 2006)	15 items, 3min	LAPS II
Fluid intelligence	CFT 20-R: Topological Deductions (Conditions) (Weiß 2006)	11 items, 3min	LAPS II
Crystallised intelligence	CFT 20-R: Number Sequences (Weiß 2006)	21 items, 12min	LAPS I
Visual working memory	Corsi Blocks	LAPS I: Start with 2 squares, 2 trials per level, 1 out of 2 trials must be correct to reach next level. LAPS II: Start with 2 squares, 3 trials per level, 1/3 trials must be correct to reach next level.	LAPS I&II
Verbal working memory	Digit Span (Forward/Backward)	LAPS I: Start with 3 digits, 2 trials per level, 1/2 trials must be correct to reach next level. LAPS II: Start with 2 digits, 3 trials per level, 1/3 trials must be correct to reach next level.	LAPS I&II
Automatic letter access, retrieval, and production	Alphabet Task (Berninger et al. 1992)	Time limit: 60s, Scoring: number of legible letters in the correct alphabetic order in the first 15s	LAPS II
Creativity (divergent thinking)	Test of creative thinking (divergent production) (TCT-DP) (Urban & Jellen 1995)	Maximum score: 72 no time constraints	LAPS I
Field independence	Group embedded figures test (GEFT) (Witkin et al. 2014)	Part 1: 7 training items, 2min, not scored; Part 2: 9 test items, 5min, Part 3: 9 test items, 5min Total score: 18	LAPS I&II

2 Language Aptitude at Primary School (LAPS): Research design

Table 4: Description of language proficiency tests

Subdimension	Test	Conditions	Study
School language German reading comprehension	ELFE 1–6, Reading Proficiency (Lenhard & Schneider 2006)	72 words, 28 sentences, 20 short texts Time limits: LAPS I 4 th /5 th grade: 3/2min _{word} , 3/2min _{sentence} , 7/6min _{text} LAPS II reduced limits at T2 and T3 (cf. technical report)	LAPS I&II
English listening and language use	Oxford Young Learners Placement Test (Oxford English Testing 2013)	language use: 18 items Listening: 12 items, duration: no time constraints, approx. 35min	LAPS I (T2)&II (T1)
General foreign language proficiency	C-tests French	4 independent texts, 20 gaps per text. Maximum score: 80 5min per text (20min)	LAPS I
General foreign language proficiency	C-tests English	5 independent texts, 20 gaps per text. Maximum score: 100 4min per text (20min)	LAPS II (T2, T3)

We modelled the C-tests on a version developed by Porsch & Wilden (2017) for a similar target group of young learners in Germany. We also consulted C-tests for teenage learners of English by Babaii & Shahri (2010) to be able to capture higher levels of competence, i.e. to avoid ceiling effects. To make sure that the texts would be appropriate in terms of vocabulary knowledge, we consulted the English manuals used in the LAPS II region to identify content areas. We adapted the texts to include topics the children were likely to be familiar with. The C-tests were piloted with three classes ($1 \times 4^{\text{th}}$ grade, $2 \times 5^{\text{th}}$ grades, $2 \times 6^{\text{th}}$ grades) who did not participate in LAPS II.

3.6.2 Other subdimensions

Modifications to the test battery were also made for intelligence, phonetic coding ability, working memory, and creativity.

In LAPS I, we used a measure of crystallized intelligence (number sequencing) to account for cognitive abilities unrelated to language. This was substituted with a test of fluid intelligence (matrices). Fluid intelligence was judged to be a more accurate assessment of general learning abilities, as it is independent of academic knowledge, such as reflected in the number sequencing test. We chose two subtests from a culture fair test (CFT 20-R, Weiß 2006) with language-free and descriptive test items.

Originally, we assessed the aptitude subcomponent of phonetic coding ability with the LLAMA-E (Meara et al. 2001). The LLAMA-E is a measure of sound-symbol association and phonemic working memory. After discussions with our panel of experts, the sound-symbol aspect of this test was deemed too closely related to literacy skills, rather than the phonemic part of language aptitude which we intended to target. In order to have a more robust indication of the phonemic aptitude component, we therefore opted for the LLAMA-D subtest (Meara et al. 2001) which measures phonemic discrimination and phonemic memory.

We also decided to strengthen the working memory (WM) measure by complementing the forward digit span for verbal WM with a backward version, which some authors argue is also a measure of the central executive component (for an overview see e.g. St Clair-Thompson & Allen 2013, Hilbert et al. 2014).

The Berninger-Graham Alphabet Task (Berninger et al. 1992) was added as a speed test for automatic letter access, retrieval, and production. The Alphabet Task is easy to administer and has been found to be predictive of children's levels of composition in the L1 (Berninger et al. 1997, Graham et al. 2006). The test was chosen with regard to our aim to explore robust predictors for L2 proficiency. If the Alphabet Task turned out to be among them, it would be a convenient option for teachers wishing to assess their students' L2 potential. To our knowledge, this possibility has not been explored previously.

We tried to accommodate these changes without adding to test taking time. We therefore decided to omit the TCT-DP for creative thinking from the test battery. This choice seemed justified, as creativity in connection with the task-based learning environment did not yield strong associations with learning outcomes in LAPS I (see Chapter 6).

4 Design

This section details participant characteristics, recruitment and procedures adopted in the LAPS I and LAPS II subprojects. In terms of the total number of participants, we draw attention to the fact that Tables 5 and 8 refer to the entire sample. The numbers reported in subsequent chapters may vary, e.g. due to the exclusion of students with L1 English or French.

The test battery was administered by members of the LAPS team and/or trained research assistants. All instructions and time limits were recorded and played via speakers to have maximum control over the elicitation process, i.e. to create situations that were as similar as possible.

4.1 Recruitment

Participants for LAPS I and LAPS II were recruited via school administrators. Teachers decided voluntarily if they wanted to participate with their class. Written consent was obtained from the pupils' parents.

4.2 LAPS I

In spring 2017, the test battery presented in Tables 2.1 to 2.4 was administered to 174 primary school pupils from 10 different classes from nine different schools (T1). In spring 2018, a second data collection (T2) took place in order to investigate students' L3 (English) proficiency as well as French and English learning motivation (see Berthele & Udry 2019 for an analysis of the skills in both FLs). Nine out of 10 classes from T1 participated in the second data collection. Table 5 presents an overview of the participants of LAPS I.

Note that the variable for multilingualism is binary and based on the parent questionnaire administered at T1. Children being classified as multilingual met at least one of the following criteria:

- L1 other than German
- Family language currently used: Other than German or German plus other
- Literacy skills in a home language other than German
- Child has received or is receiving tuition for non-native German speakers (referred to in the Swiss system as *Deutsch als Zweitsprache*, DaZ)
- Child has received or is receiving tuition in the home language (referred to in Switzerland as *Heimatliche Sprache und Kultur*, HSK)

The number of L2 and L3 lessons is an estimated average of the total tuition received at the time of testing. The estimate is based on an evaluation of the cantons' current timetables which was mandated by the government and conducted by Bucher & Zemp (2019). According to the authors, the L2 is taught to Swiss primary school children with a total of 390 45-minute lessons. These lessons are distributed over 4 academic years (comprising 39 weeks each), usually with three weekly lessons in grades 3 and 4, and two weekly lessons in grades 5 and 6.

Depending on the canton, the number of L3 lessons taught at primary school ranges from 152 to 234.³ The total varies because of different timetables in grades 5 and 6, the period of L3 instruction: Some cantons opt for three weekly lessons per grade, while others offer only two weekly lessons per grade.

Table 5: Participants LAPS I

Time	Grade	N	Multiling.	Age		Mean <i>n</i> lessons	
				Mean	Range	L2 French	L3 English
T1	4	57	11	10.3	9.4–11.0	172	0
T2	5	55	10	11.3	10.5–11.6	248	58
T1	5	117	27	11.5	10.6–13.2	248	58
T2	6	103	23	12.6	10.5–14.2	324	134
T1		174	38	11.1	9.4–13.2		
T2		158	33	12.1	10.5–14.2		

4.3 LAPS I procedure T1

Data collection took place between March and April 2017. Tests were administered by two or three assistants or LAPS-researchers, depending on the number of students per class. The administration of the entire test battery took approximately 3 hours and was divided into two sessions in order to prevent fatigue scheduled within a week (see Table 6). The test sequence was organized to allow for alternation between more and less cognitively demanding tasks. For practical reasons the order of the tasks could not be varied. It is possible that there are order effects of the tasks. However, if so, they arguably affect the participants in a similar way. Since we are interested in differences among the students, we deem the impact of such order effects on the overall results to be modest.

³In 2019, Appenzell Innerrhoden was the only canton to introduce L3 teaching at secondary school.

Table 6: Procedure LAPS I, T1

Session 1 (≈ 100 min)	Slot 1	Introduction; Language Detective; GEFT
	Slot 2	Team up Words!; ELFE 1–6
Session 2 (≈ 90 min)	Slot 3	C-test French; TCT-DP; CFT 20-R (numerical sequences)
	Slot 4	LLAMA-E; Corsi Blocks; Digit Span; Motivation Questionnaire

4.4 LAPS I procedure T2

Data collection took place a year later in April and May 2018. We administered measures of L3 English proficiency and motivation in one session that lasted approximately 75 minutes. This session was conducted by two assistants or LAPS researchers, applying standardized instructions and procedures. Language tests and questionnaires were alternated (see Table 7).

Table 7: Procedure LAPS I T2 – Spring 2018

Test session ^a	Introduction; Questionnaire L2 English; Oxford Young Learners Placement Test; Questionnaire L3 French
---------------------------	---

^a(Short breaks between tasks)

4.5 LAPS II

After LAPS I, minor changes discussed in section 3 were made to the test battery. In autumn 2017 (T1), the adapted version was administered to primary school pupils from 32 different classes (13 4th graders, 15 5th graders and 4 mixed grade classes). As in LAPS I, instructions and procedures were standardized and introduced to research assistants in a training session.

In spring 2018 (T2) and 2019 (T3) five measures were re-administered to the same participants to monitor longitudinal development: L2 English proficiency, English/French motivation questionnaire, language of instruction German, grammatical sensitivity, and inductive ability. A total of 637 pupils participated in LAPS II. Table 8 provides participant details. Note that the same criteria apply to the variable *multilingual* as in LAPS I.

Table 8: Participants LAPS II. T1: Autumn 2017; T2: Spring 2018; T3: Spring 2019

Time	Grade	N	Multiling.	Age		Mean <i>n</i> lessons	
				Mean	Range	L3 French	L2 English
4 th graders at T1							
T1	4	289	171	10.0	9.1–11.6		181
T2	4	274	161	10.6	9.7–12.2		259
T3	5	260	150	11.6	10.7–13.2	66	346
5 th graders at T1							
T1	5	326	157	11.0	9.2–12.4	14	294
T2	5	304	143	11.6	9.8–13.0	66	346
T3	6	306	146	12.6	10.8–14.0	146	426
Total participants							
T1	615	328	10.5	9.1–12.4			
T2	578	304	11.1	9.7–13.0			
T3	566	296	12.1	10.7–14.0			

4.6 LAPS II procedure T1

The administration of the entire test battery took approximately 3.5 hours. Similar to LAPS I, it was divided into two sessions (see Table 9). Nine assistants were recruited and trained for test administration.

4.7 LAPS II procedure T2 and T3

At T2 and T3, selected measures from the test battery were repeated to longitudinally track their development: L2 English proficiency, language of instruction German, motivation, grammatical sensitivity, and inductive ability. They were administered in one session of approximately 90 minutes divided into two slots (Table 10). As only paper and pencil group tests were administered at T2 and T3, fewer assistants were needed for data collection. Four research assistants at T2 and two research assistants at T3 were recruited and trained analogous to T1.

Table 9: Procedure LAPS II T1: Autumn 2017

Session 1 (≈ 120 min)	Slot 1	Introduction; Language Detective; GEFT
	Slot 2	ELFE 1–6; CFT 20-R
	Slot 3	Alphabet Task; Team up Words!
Session 2 (≈ 90 min)	Group A 2 A	LLAMA-D; Corsi blocks; Digit Span (f/b); Motivation Questionnaire English
	Group B 1A+ teacher	Oxford Young Learners Placement Test (OYLPT); Motivation Questionnaire English

Table 10: Procedure LAPS II T2 and T3: Spring 2018 & 2019

Test session 1–2TL	Slot 1	Introduction; Language Detective Team up Words!; English Motivation Questionnaire
	Slot 2	ELFE 1–6; English Proficiency (C-test); French Motivation Questionnaire

5 Scoring and data entry

In LAPS I and LAPS II, computer-administered tests were scored automatically. Paper and pencil tasks were scored by members of the LAPS team and assistants who participated in the data collection. Subsequently, all data were entered and stored for analysis. To minimise the chances of mishaps in data entry, a special platform was created restricting the data format of the input and displaying error messages when impossible data was entered (cf. Vanhove 2018).

Test scoring and data entry procedures were defined in a manual and the research assistants were trained accordingly. The most important aspects of the process will be summarised in the following paragraphs. Full details on handling missing values and scoring the tests are given in the technical report (<https://osf.io/d9gnh/>).

For the English proficiency tests administered at T2 and T3, two different scoring types were applied:

- Spelling errors are penalized: This constitutes the standard scoring type.
- Spelling errors are ignored: This accounts for the communicative language learning approach practised in Swiss EFL-classrooms which doesn't focus strongly on spelling.

Due to the acceptance of phonetically correct spellings, the second scoring type opens the door to variable criteria as to what represents an acceptable response and what not. When trialling the C-tests, lists of accepted and unaccepted spellings were defined. At T2 and T3, a randomly selected subset was scored by two independent raters and the guidelines for spelling variants were supplemented accordingly.

A subsequent analysis reveals that the two scoring types are strongly correlated ($r_{T2} = 0.98$, $r_{T3} = 0.98$, cf. technical report, Figure 16.1 and Figure 17.1). For any statistical analysis we used the first scoring type, i.e. where spelling errors are penalized.

References

- Babaii, Esmat & Somaiyeh Shahri. 2010. Psychometric rivalry: The c-test and the cloze test interacting with test takers' characteristics. In Rüdiger Grotjahn (ed.), *Der C-Test: Beiträge aus der aktuellen Forschung/ The C-Test: Contributions from Current Research*, 41–56. Frankfurt am Main: Lang.
- Berninger, Virginia W., Katherine B. Vaughan, Robert D. Abbott, Sylvia P. Abbott, Laura Woodruff Rogan, Allison Brooks, Elizabeth Reed & Steve Graham. 1997. Treatment of handwriting problems in beginning writers: Transfer from handwriting to composition. *Journal of Educational Psychology* 89(4). 652–666.
- Berninger, Virginia W., Cheryl Yates, Ana Cartwright, Judith Rutberg, Elizabeth Remy & Robert Abbott. 1992. Lower-level developmental skills in beginning writing. *Reading and Writing* 4(3). 257–280. DOI: 10.1007/BF01027151.
- Berthele, Raphael & Isabelle Udry. 2019. Multilingual boost vs. cognitive abilities: Testing two theories of multilingual language learning in a primary school context. *International Journal of Multilingualism*. 1–20. DOI: 10.1080/14790718.2019.1632315.
- Bildungsdirektion Kanton Zürich. 2017. *Lehrplan 21: Sprachen*. Zürich: Lehrmittelverlag Zürich.

- Bucher, Monika & Benedict Zemp. 2019. *Studentafeln zum Lehrplan 21: Auswertung 2017/2018/2019*. Luzern: BKZ Geschäftsstelle Luzern.
- Carroll, John B. 1958. A factor analysis of two foreign language aptitude batteries. *The Journal of General Psychology* 58(1). 3–19. DOI: 10.1080/00221309.1958.9710168.
- Carroll, John B. & Stanley M. Sapon. 2010. *Modern language aptitude test: Elementary manual*. Rockville: Second Language Testing Foundation.
- Dörnyei, Zoltán. 2010. *Questionnaires in second language research*. 2nd edn. New York: Routledge.
- Eckes, Thomas & Rüdiger Grotjahn. 2006. A closer look at the construct validity of c-tests. *Language Testing* 23(3). 290–325. DOI: 10.1191/0265532206lt330oa.
- EDK. 2004. *Sprachenunterricht in der obligatorischen Schule: Strategie der EDK und Arbeitsplan für die gesamtschweizerische Koordination*. Bern: Schweizerische Konferenz der Kantonalen Erziehungsdirektoren.
- Ellis, Rod. 2017. Task-based language teaching. In Shawn Loewen & Masatoshi Sato (eds.), *The Routledge handbook of instructed second language acquisition*, 108–125. New York: Routledge.
- Graham, Steve, Miriam Struck, Julie Santoro & Virginia W. Berninger. 2006. Dimensions of good and poor handwriting legibility in first and second graders: Motor programs, visual–spatial arrangement, and letter formation parameter setting. *Developmental Neuropsychology* 29(1). 43–60. DOI: 10.1207/s15326942dn2901_4.
- Heinzmann, Sybille. 2013. *Young language learners' motivation and attitudes: Longitudinal, comparative and explanatory perspectives*. London: Bloomsbury.
- Hilbert, Sven, Tristan T. Nakagawa, Patricia Puci, Alexandra Zech & Markus Bühner. 2014. The digit span backwards task. *European Journal of Psychological Assessment* 31(3). 174–180. DOI: 10.1027/1015-5759/a000223.
- Horwitz, Elaine, Michael B. Horwitz & Joann Cope. 1986. Foreign language classroom anxiety. *The Modern Language Journal* 70(2). 125–132.
- Lenhard, Wolfgang & Wolfgang Schneider. 2006. *ELFE 1-6 Ein Leseverständnistest für Erst- bis Sechstklässler*. Göttingen: Hogrefe.
- Meara, Paul M., James Milton & Nuria Lorenzo-Dus. 2001. *Llama language aptitude tests: The manual*. Swansea.
- Nowicki, Stephen & Bonnie R. Strickland. 1973. A locus of control scale for children. *Journal of Consulting and Clinical Psychology* 40(1). 148–154.
- Oxford English Testing. 2013. *Oxford young learners placement test*. <https://www.oxfordenglishtesting.com/> (5 May, 2017).

- Passepartout, Arbeitsgruppe Rahmenbedingungen. 2008. *Passepartout: Didaktische Grundsätze*. <http://www.passepartout-sprachen.ch/informationen-fuer/lehrpersonen-primarstufe/worum-geht-es/> (5 July, 2017).
- Peyer, Elisabeth, Mirjam Andexlinger, Karolina Kofler & Peter Lenz. 2016. *Projekt Fremdsprachenevaluation BKZ: Schlussbericht zu den Sprachkompetenztests*. Tech. rep. Université de Fribourg. <http://doc.rero.ch/record/259287/> (9 February, 2018).
- Pimsleur, Paul, Daniel J. Reed & Charles W. Stansfield. 2004. *Pimsleur language aptitude battery PLAB*. Rockville: Second Language Testing.
- Porsch, Raphaela & Eva Wilden. 2017. The development of a curriculum-based c-test for young EFL learners. In J. Enver & E. Lindgren (eds.), *Researching the complexity of early language learning in instructed context* (Multilingual Matters), 289–304. Clevedon: Multilingual Matters.
- St Clair-Thompson, Helen L. & Richard J. Allen. 2013. Are forward and backward recall the same? A dual-task study of digit recall. *Memory & Cognition* 41(4). 519–532. DOI: 10.3758/s13421-012-0277-2.
- Stöckli, Georg. 2004. *Motivation im Fremdsprachenunterricht: Eine theoriegeleitete empirische Untersuchung in 5. und 6. Primarschulklassen mit Unterricht in Englisch und Französisch* (Pädagogik bei Sauerländer). Oberentfelden: Sauerländer Verlag.
- Urban, Klaus K. & Hans G. Jellen. 1995. *Test zum Schöpferischen Denken – Zeichnerisch*. Frankfurt am Main: Pearson PsychCorp.
- Vanhove, Jan. 2018. *A data entry form with sanity checks*. <https://janhove.github.io/design/2018/07/06/data-entry-failsafes> (31 July, 2020).
- Weiß, Rudolf H. 2006. *Manual Grundintelligenztest Skala 2 – Revision (CFT 20-R)*. Göttingen: Hogrefe.
- Witkin, Herman A., Philip K. Oltman, Evelyn Raskin, Stephen A. Karp & Jack Demick. 2014. *Group embedded figures test manual*. Menlo Park: Mind Garden.

Chapter 3

The smart, the motivated and the self-confident: The role of language aptitude, cognition, and affective variables in early instructed foreign language learning

Isabelle Udry^{a,b} & Raphael Berthele^a

^aUniversity of Fribourg, Institut de Plurilinguisme ^bZurich University of Teacher Education

We investigated the underlying structure of language-related, cognitive, and affective variables assessed in a test battery. First, we conducted an exploratory factor analysis (EFA) with data from LAPS I (174 learners of L2 French, mean age 11.1). Second, we conducted a confirmatory factor analysis (CFA) with data from LAPS II (615 learners of L2 English, mean age 10.5). EFA with LAPS I data yielded an optimal solution with three dimensions: (1) General cognitive abilities and language aptitude (Cognition/Aptitude Factor); (2) extrinsic motivation, dedication, parental and teacher encouragement (Extrinsic Factor); (3) intrinsic motivation, foreign language anxiety, and L2 self-concept (L2 Academic Emotion Factor). The CFA with sample LAPS II validated the EFA factor structure. The findings provide robust evidence for a cognitive dimension involving language aptitude, IQ, working memory, field independence and language of instruction, even when context and L2 differ.

A regression analysis was carried out with the 3 factors as predictors and L2 proficiency as the dependent variable. It reveals that Factor 1 (Cognition/Aptitude) and Factor 3 (L2 Academic Emotion) are positively associated with L2 proficiency. When the two other factors are controlled for, a negative effect of the Extrinsic Factor on L2 proficiency is observed in the data.

Isabelle Udry & Raphael Berthele. 2021. The smart, the motivated and the self-confident: The role of language aptitude, cognition, and affective variables in early instructed foreign language learning. In Raphael Berthele & Isabelle Udry (eds.), *Individual differences in early instructed language learning: The role of language aptitude, cognition, and motivation*, 71–89. Berlin: Language Science Press. DOI: 10.5281/zenodo.5464749



1 Introduction

The goal of this chapter is to explore the relationship between individual differences (IDs) and their impact on L2 learning outcomes. To this aim, we administered a test battery of language-related, cognitive, and affective variables to primary school children aged 10–12 years, learning foreign languages as part of the mandatory curriculum. The conceptual framework of these variables is discussed in Chapter 1, while a full description of the project Language Aptitude at Primary School (LAPS) and the test battery can be found in Chapter 2. The following research questions were addressed:

1. What is the underlying structure of the ID variables?
2. Is the structure replicable in an independent sample?
3. What are the effects of the ID variables on L2 proficiency?

Chapter 4 deals with predictive models involving all IDs and environmental factors assessed in the LAPS project, and Chapter 5 considers the impact of environmental factors. This chapter is concerned with the latent constructs relating to the individual only and their relationship with instructed foreign language learning.

2 Design and Procedure

To describe patterns of IDs as they emerge from the data, we adopted a data-driven approach based on recommendations by Brown (2006). The factor analysis was conducted in two steps: In step 1, we performed an exploratory factor analysis (EFA) on the sample from LAPS I (first data collection T1, L2 learners of French). In step 2, we applied confirmatory factor analysis (CFA) to the data from LAPS II (first data collection T1, L2 learners of English). Finally, a regression analysis was carried out with both samples with L2 proficiency as the dependent variable. This procedure allowed for a maximally unbiased approach to the data.

As one of the reviewers rightly pointed out, other statistical procedures involve testing pre-specified, theory-based models with the same data and then deciding on the most appropriate model, according to goodness-of-fit indices and theoretical assumptions. Since we did not intend to test a specific theory of language learning, this approach seemed less suitable for our context. Moreover, as outlined in Chapter 4, it is not uncommon for several competing models to yield an acceptable goodness-of-fit. Statistical analysis that starts with applying

different theory-based models to the same sample, to our mind, bears the risk of circular thinking whereby a theory that has not been fully validated (and that the study may intend to test) is eventually used to corroborate this theory, even though other models may also have performed within an acceptable range.¹ We therefore chose the procedure described above which we deemed to be most appropriate for our descriptive approach and our research questions.

2.1 Sample

Native speakers of French and English were excluded from the samples. Since these participants cannot be considered L2 learners of the respective languages, their answers in the questionnaire assessing motivation for L2 learning were expected to be incongruent with those of L2 learners. Incomplete data points were also removed from the data. This resulted in the following datasets:

The exploratory factor analysis was conducted with data from LAPS I which consists of 174 L2 French learners, mean age 11.1 years (4th and 5th grade) with L1 German. The data were collected in 2017 in 10 classes located close to the French-speaking region of Switzerland. These children started to learn L2 French in 3rd grade, aged 9.

The confirmatory factor analysis was conducted with data from LAPS II and consists of 615 L2 English learners, mean age 10.5 years (4th and 5th grade) with L1 German. Data collection took place in autumn 2017 in the Swiss German speaking part of the country in 32 classes. Children from this sample had already started with L2 English lessons in grade 2.

2.2 Test instruments

A test battery including measures of language aptitude, cognition, and affective variables was compiled and administered on four lessons of 45 minutes distributed over two sessions on different days.²

2.2.1 Language aptitude

Inductive ability (ind): Participants are presented with words and short sentences in an artificial language, as well as their translation in the school

¹For a discussion, see for instance Vafaei et al. (2017) on modelling measures of explicit versus implicit L2 knowledge.

²Full details on the test battery are given in Chapter 2.

language. The participants' task is to infer regularities and translate sentences following the same pattern from the school language into the artificial language. This is inspired by a task (Form 4) of the Pimsleur language aptitude battery (PLAB, Pimsleur 1966). Task design (structure of the artificial language and instructions) was customized to fit our target group.

Grammatical sensitivity (gra): Participants are presented with two sentences. In the first one, a word is highlighted. The participants' task is to find, in the second sentence, the word having the same function as the highlighted word in the first sentence. This task is based on Part 2 ("Matching Words") of the Modern Language Aptitude Test - Elementary (MLAT-E, Carroll & Sapon 1976) and was presented in the language of instruction: German.

Phonetic coding ability (pho): We used two subtests of Meara et al.'s (2001) Llama test battery.

- In the LLAMA E (sound-symbol association, phonemic working memory), participants have two minutes to learn how sounds are spelled in an artificial language (training phase). Their task is then to listen to bi-syllabic words in the artificial language and pick their correct spelling from among two options (used in LAPS I).
- LLAMA D (sound recognition task, phonemic discrimination): In a trial phase, participants hear different sounds. During testing, they listen to strings of sounds and are asked to identify the sounds from the trial phase (used in LAPS II).

2.2.2 General learning abilities

Intelligence (iq): CFT 20-R (Weiß 2006). The participants complete number sequences (LAPS I) and matrices/topological deduction (LAPS II).

Field (in-)dependence (fie): In the Group Embedded Figures test (GEFT, Witkin et al. 2014), participants have to find simple geometrical figures embedded in more complex figures under time pressure.

Working memory:

- *Visuospatial working memory (vis)*: We administered an adaptation of the Corsi Block Task, in which participants need to remember the order of an increasing number of squares lit up from a matrix of squares.

- *Verbal working memory (vem)*: A Forward Digit Span task, in which participants are asked to reproduce series of numbers of increasing length. Stimuli are presented both visually and aurally.

2.2.3 Affective variables

A student questionnaire assessing language learning motivation was put together based on previous work by Horwitz et al. (1986), Stöckli (2004), Dörnyei (2010), Heinzmann (2013), and Peyer et al. (2016). It included the following sub-dimensions: intrinsic motivation, extrinsic motivation (school/leisure), lingua franca motivation, foreign language learning anxiety, self-concepts (L2 + school language German), teacher motivation, parental encouragement, dedication, and future L2 self. In LAPS I, the focus was on L2 French and in LAPS II on L2 English. The wording of the items remained unchanged except for the language label.

In addition, Locus of Control (loc) was measured with a translation of the N-S Personality scale (based on Nowicki & Strickland 1973).

2.2.4 Language proficiency

Proficiency in the language of instruction – German (ger) – was measured with ELFE (Lenhard & Schneider 2006), a standardized reading comprehension test at word, sentence and text level.

L2 proficiency was the dependent variable in the regression analysis. It was assessed with

- C-Tests for general L2 proficiency in LAPS I (L2 French)
- Oxford Young Learners Placement Text (OYLPT) (Oxford English Testing 2013) in LAPS II (L2 English). The OYLPT measures oral and written comprehension embedded in everyday, communicative situations.

2.2.5 Adaptations

The same constructs were assessed in samples 1 and 2. However, due to the overall logic of the LAPS project, two changes were made with regards to the tests administered for phonetic coding ability and intelligence (cf. Chapter 2, §3.6). After discussion with a panel of experts, the LLAMA E (sound-symbol association) was replaced by the LLAMA D test (phonemic discrimination, cf. Meara et al. 2001) for LAPS II. LLAMA D was deemed more appropriate to assess the phonological aspect of language learning targeted in the study. The LLAMA E is similar to grapheme–phoneme correspondence and therefore to reading skills.

The IQ test involving the completion of number sequences used in sample 1 was replaced by a module of the same intelligence test battery CFT (Weiß 2006) that tests the ability to complete graphical matrices (fluid intelligence). This change was also recommended by the same panel of experts. Fluid intelligence was judged to be more suitable because of its independence from academic knowledge which may have been a limitation of the number sequencing test.

3 Factor Analysis

The purpose of factor analysis is to uncover the smallest number of latent variables (factors) underlying a set of observed variables (Brown 2006). Exploratory factor analysis (EFA) makes no prior assumptions about the pattern of relationships between the observed measures and the latent variables (Brown 2006). Confirmatory factor analysis (CFA), on the other hand, aims to replicate structures that have been found previously. Therefore, several elements of the CFA model, such as the number of factors, are specified in advance.

In the LAPS project, the procedure was applied as follows: (1) Choose the appropriate estimator and estimate the factor model; (2) select the appropriate number of factors; (3) select a rotation technique to obtain simple structure in order to interpret the factors; (4) replicate the analysis with an independent sample, i.e. with confirmatory factor analysis (Brown 2006). A more technically detailed description of the methods and procedures used in this chapter are provided as supplementary online material at <https://osf.io/tpshc/>.

3.1 Exploratory Factor Analysis (LAPS I, L2 French)

We employed the `fa()` function from the Psych-package in R (William 2018) using a maximum likelihood method. The latent variables in this data are assumed to correlate, therefore we used an oblique rotation with promax (Bortz & Schuster 2010: 419). Based on the common methods for factor selection, a 3-factor solution was chosen.

Factor loadings > 0.3 were considered to be meaningful for interpretation (Brown 2006: 30). This yields the following structure (Table 3.1): Factor 1 is associated with variables of general learning abilities (IQ, WM, field independence), as well as aptitude and language-related abilities (grammatical sensitivity, inductive ability, phonetic coding, and L1 skills). We suggest the label *Cognition/Aptitude* for this factor. It accounts for 16% of the variance. Factor 2 subsumes variables linked to external influences, such as teacher and parental encouragement, extrinsic motivation and dedication. This *Extrinsic Factor* accounts for 12% of the

variance. Factor 3 includes intrinsic motivation, foreign language anxiety and L2 self-concept. We refer to the third factor by the label *L2 Academic Emotion*. It accounts for 12% of the variance. The model explained 40% of the total variance.

Two variables could not be clearly tied in to only one of the three factors. Intrinsic motivation was associated most strongly with L2 Academic Emotion (0.77), a finding that is congruent with the literature (Csizér & Kormos 2009, Liu & Huang 2011, Noels et al. 2000, Stöckli 2004). Intrinsic motivation also loaded moderately on the Extrinsic Factor 2, indicating that the two types of motivation are related (see discussion in §3.3). Second, locus of control yielded modest loadings on the Cognition/Aptitude Factor 1 (−0.32) and L2 Academic Emotion Factor 3 (−0.30). We used different analytical scenarios in the CFA to account for and clarify the ambiguous role of intrinsic motivation and locus of control found in this exploratory analysis.

Table 1: Loadings of the three-factor solution. Loadings with an absolute value of > 0.3 are in bold. Factors are: 1 – Cognition/Aptitude, 2 – Extrinsic, 3 – L2 Academic Emotion.

Variables	Factors		
	1	2	3
intelligence (iq)	0.68	−0.10	0.00
field independence (fie)	0.66	−0.05	−0.06
grammatical sensitivity (gra)	0.62	0.13	0.03
verbal WM (vem)	0.57	0.05	−0.16
phonetic coding (pho)	0.51	0.09	−0.04
visual WM (vis)	0.50	−0.06	−0.06
school language German proficiency (ger)	0.46	−0.07	0.12
inductive ability (ind)	0.44	0.00	0.07
locus of control (loc)	−0.32	−0.04	−0.31
dedication (ded)	0.00	0.69	0.13
extrinsic motivation (ext)	0.01	0.67	−0.11
teacher encouragement (tea)	0.04	0.65	0.11
parental encouragement (par)	−0.01	0.56	−0.12
intrinsic motivation (int)	−0.11	0.36	0.70
L2 self-concept (sel)	−0.12	0.03	0.77
anxiety (anx)	−0.05	0.26	−0.77

3.2 Confirmatory Factor Analysis: LAPS II (L2 English)

The second part of our analyses was based on the factor structure yielded by the EFA. Our aim was to test whether the patterns from EFA could be found in other populations of schoolchildren learning a different target language. To do this, we used data from LAPS II. The test battery was largely identical to the battery used in sample 1 (L2 French), with some minor modifications discussed in §2.2.5.

Confirmatory factor analysis puts a factorial structure found in an exploratory approach to the test. It follows the logic of fitting a structural equation model (SEM) to the data that is parametrized based on these previous findings. The factors from the EFA thus represent the latent constructs in the SEM model of the CFA. Based on suggestions in van Prooijen & van der Kloot (2001), we fitted several model variants. The loadings found in the EFA were used to determine the associations of variables to factors. The differences in associations of variables to the three factors across the four models listed in Table 2 help clarify the ambiguous status of the variables locus of control and intrinsic motivation discussed above. These models and their goodness-of-fit indices are given in Table 2; more details as well as an additional model fitting the exact loadings of the EFA to the new data can be found in the supplementary material.

Table 2: Models tested in the CFA with goodness-of-fit indices. IM: Intrinsic motivation on factor(s) (numbers). LC: Locus of control on factor(s) (numbers).

Variant	Path		Goodness-of-fit			
	IM	LC	χ^2	CFI	RMSEA	SRMR
1 ^a	3	1	407.04	0.87	0.07	0.07
2 ^b	2 + 3	1 + 3	314.92	0.91	0.06	0.06
3 ^c	3	excluded	306.59	0.9	0.06	0.07
4 ^d	2 + 3	excluded	223.56	0.94	0.05	0.06

^aConsider only highest loading.
^bAllow cross-loadings (i.e. variables with two loadings >0.3 load onto two factors).
^cConsider highest loadings, exclude locus of control.
^dAllow loadings on two factors, exclude locus of control.

We fitted all models with the `cfa()` function of the `lavaan` package (Rosseel 2012) with the MLR estimator, which is recommended (Hallquist 2018) for data that are not all perfectly normally distributed, clustered, and may contain missing values.

The following indices are generally reported for SEM (Kline 2011), with suggested criteria for the assessment of the model fit provided in brackets: χ^2 ; root mean square error of approximation RMSEA (< 0.08); comparative fit index CFI (> 0.9); standardized root mean residual SRMR (< 0.08). The four models yield an acceptable fit to the data (Table 2). The goodness-of-fit indices improve slightly if locus of control is excluded altogether from the analysis and intrinsic motivation is allowed to load on the two affective factors (variant 4). Our findings suggest that the underlying factorial structure identified in the EFA can indeed be found again in the CFA, even though the target language is a different one. Figure 1 shows the path diagram of these factors (variant 4), i.e. Cognition/Aptitude, L2 Academic Emotion, and Extrinsic.

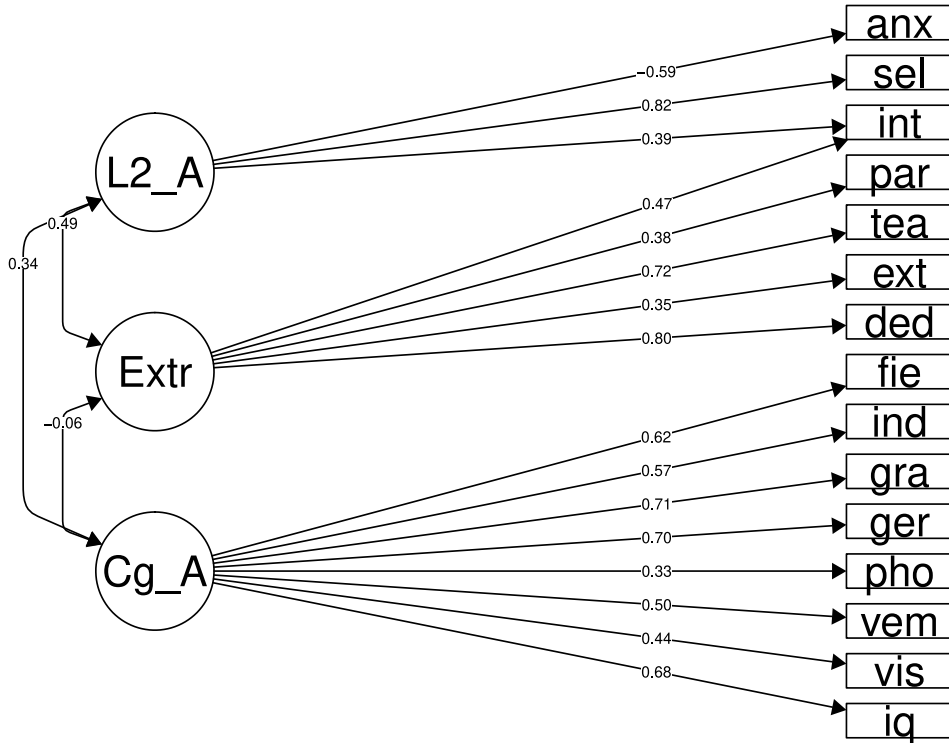


Figure 1: Confirmatory factor analysis of the three-factor solution, variant 4; (L2_A: L2 Academic Emotion; Extr: Extrinsic; Cg_A: Cognition/Aptitude)

3.3 Discussion of the factor analysis

EFA was conducted with a sample of 174 L2 French learners (mean age 11.1). We then applied the structure to an independent sample of 615 L2 learners of English (mean age 10.5) by means of CFA to find out if the same structure could be uncovered in a different context and with a different target language. The EFA yielded three factors that were validated by the CFA: 1) Aptitude/Cognition, 2) Extrinsic Motivation, 3) L2 Academic Emotion. We will discuss these factors in turn.

The first factor (Aptitude/Cognition) regroups all cognitive and language-related ID variables. It explains most of the variance in the EFA model (16%). The result suggests that for the young learners in this study, language-related and other cognitive variables are not clearly separable, but represent a common latent factor. Our data do not corroborate the presence of a language-specific dimension underlying these children's instructed L2 learning. Rather, it provides evidence for a cognitive dimension involving language aptitude (phonetic coding, grammatical sensitivity, inductive ability), working memory (visual and verbal), intelligence, proficiency in language of instruction (in this case German), and field independence. This aligns with more recent conceptions of language learning ability that emphasize the complementarity of domain-specific and domain-general aspects (see e.g., Skehan 2019 for a discussion). Our data suggest that the language analysis component of aptitude and general cognitive ability are associated in young learners, a link that has so far been documented mainly for older learners (Grañena 2012, Sasaki 1996). Moreover, language aptitude has been shown to be independent of, but overlapping with, intelligence in adults and adolescents (Wesche et al. 1982, Sasaki 1996, Li 2016: 827). Our findings substantiate these claims for young learners. Having said that, both intelligence and language aptitude comprise several subcomponents and some authors argue that they are associated differently (Grañena 2013, Li 2016). Since we did not assess all aspects of intelligence, we are unable to provide detailed information on the relationships between the various subcomponents.

Affective variables represent dimensions distinct from the cognitive and language-related abilities. The two affective factors explain slightly less variance (12% each). Factor 2 (labelled Extrinsic Motivation) is related to variables independent of the perceived qualities of the L2 as a school subject and includes extrinsic motivation, perceived teacher and parental encouragement, and dedication. Factor 3 (labelled L2 Academic Emotion) regroups emotions involved in L2 learning. These include enjoyment of L2 learning (intrinsic motivation), but also negative emotional states associated with it (anxiety), and children's perception of themselves as L2 learners (self-concept). The interplay of these constructs

is well documented in the literature (see Chapters 1 and 8). Learning processes in general are expected to trigger a range of emotions that mutually influence each other and ultimately impact on students' academic achievement (Pekrun et al. 2002). In terms of L2 learning, an unfavorable L2 self-concept may heighten anxiety experienced in the classroom, thus impacting negatively on a learner's ability to enjoy L2 lessons and decrease intrinsic motivation. Several authors have described this connection: L2 anxiety has been found to correlate with learners' motivational orientation in general (Heinzmann 2013: 189) and with intrinsic motivation in particular (Noels et al. 2000, Stöckli 2004, Kormos & Csizér 2008, Liu & Huang 2011). A reverse effect with poor language learning being the cause of anxiety has also been hypothesized (Sparks et al. 2011).

While the association between general cognitive abilities and aptitude variables emerged clearly from the data, there was some ambiguity associated with the affective variables *locus of control* and *intrinsic motivation* which loaded on both affective factors. The affective factors largely represent the distinction between intrinsic and extrinsic motivation from Self-Determination Theory (SDT) by Deci & Ryan (1985, 2002; see Chapter 1 for a discussion). Intrinsic motivation is a central construct of SDT, subsuming self-determination, competence, and interpersonal relatedness as three basic psychological needs. Extrinsic motivation means engagement with an activity to attain a separable outcome (Ryan & Deci 2000: 60). The two types of motivation are closely linked, i.e. they are located on a continuum, and extrinsic forms can gradually become intrinsic through the process of internalization (see e.g., Deci & Ryan 1985). For instance, an external influence, such as expected career opportunities resulting from L2 learning, may be internalized, resulting in an intrinsic wish to study the language, provided that the individual feels in charge of the learning process (i.e. the need for self-determination is met). With regard to school learning, Ryan & Deci (2000: 64) emphasize that enjoyment of and willingness to engage with contents (intrinsic motivation) is shaped by teacher and peer interaction (external influence). The connection between the intrinsic and extrinsic dimension can be seen in our data from the double affiliation of intrinsic motivation with both affective factors. Indeed, when intrinsic motivation is allowed to load on two factors, the model fit becomes slightly better.

4 Regression analysis

We conducted a regression analysis on the LAPS I and LAPS II data. Due to the similarities of the findings, only LAPS II will be reported here.³ The aim of the regression analysis was to estimate the correlation between the three factors identified in the factor analysis and participants' L2 English achievement. The dependent variable was the total score in the English test for listening and reading comprehension (OYLPT) taken at T1. Figure 2 shows the bivariate associations of the scores in the data to which the statistical model was fitted.

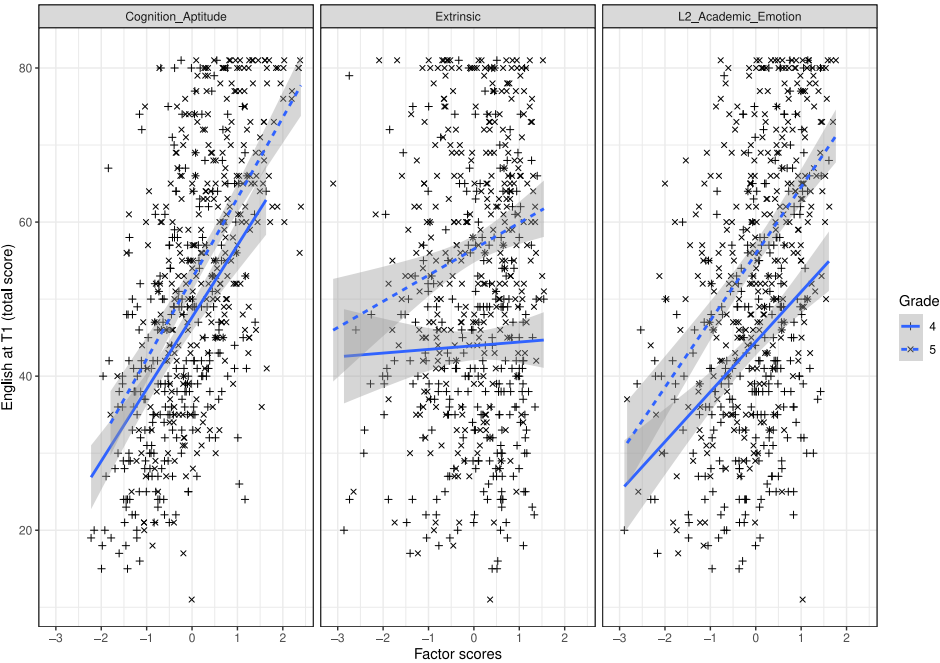


Figure 2: Bivariate associations of the factor scores and English skills at T1 for grades 4 and 5 (LAPS II data). They indicate direction and strength of the relationship between each factor and L2 proficiency.

We fitted a linear mixed model with the `lmer()` function of the `lme4` package (Bates et al. 2015). The factor scores from the CFA were introduced as independent variables (fixed effects). We included random intercepts by teacher to account for the clustered nature of the sample (random effects).

As shown in Table 3 and Figures 2 and 3, Factor 1 (cognition/aptitude) and Factor 3 (L2 Academic Emotion) are positively correlated with L2 proficiency.

³The entire analysis is available at <https://osf.io/tpshc/>.

The extrinsic Factor 2 is negatively correlated with L2 proficiency. These findings will be discussed in the next section.

Table 3: Summary of the regression model of the factor scores on English proficiency at T1

	Estimate	SE	df	<i>t</i>	Pr(> <i>t</i>)
(Intercept)	50.4	0.88	28.96	57.33	<0.001***
Cognition/apptitude	6.6	0.69	564.29	9.57	<0.001***
Extrinsic	-6.9	1.06	582.92	-6.51	<0.001***
L2 Academic Emotion	11.8	1.11	577.40	10.71	<0.001***

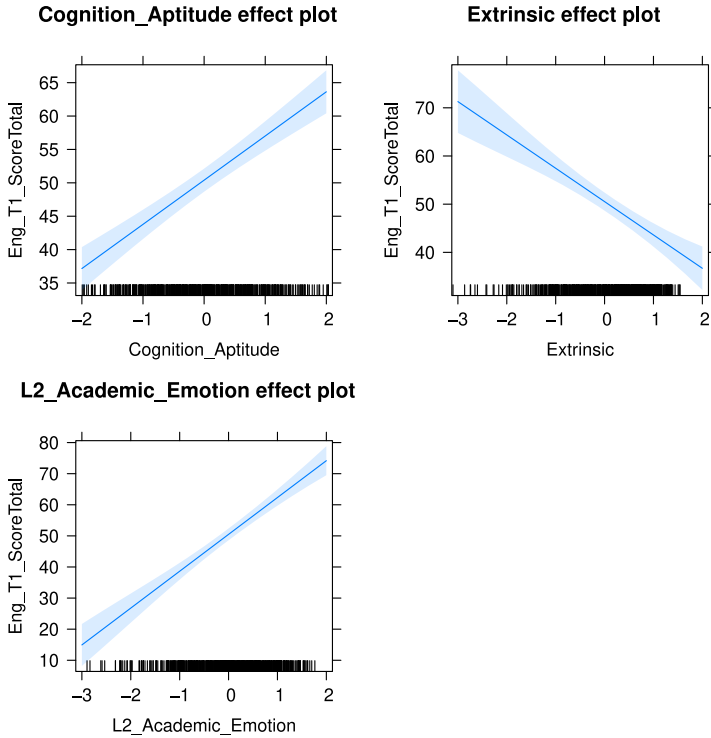


Figure 3: Partial effects of the three factor scores and English skills at T1 (LAPS II data, both grades 4 and 5)

EFA and CFA yielded a robust factor structure with three factors: 1. Cognition/Aptitude, 2. Extrinsic Motivation, 3. L2 Academic Emotion. In the final step

of the analysis, we were interested in the associations of these factors with L2 proficiency. To this aim, we considered:

1. the association between each individual factor and L2 proficiency illustrated in the bivariate correlation plots (Figure 2)
2. and the mutual relationship between the three factors and L2 proficiency in a linear mixed model (Figure 3).

Figure 2 shows that the bivariate correlations between each factor and L2 proficiency are positive. When considered together in a mixed model (Figure 3), Factor 1 (Cognition/Aptitude) and Factor 3 (L2 Academic Emotion) are positively associated with L2 outcomes while Factor 2 (Extrinsic Motivation) is related negatively to L2 proficiency. This means that what remains of a statistical effect of Factor 2 on the outcome variable after controlling for the two other, more strongly associated constructs, is a negative effect. As some reviewers pointed out, the bivariate correlations and results from the regression analysis seem to contradict each other. It is worth noting that the two stand for different things: Bivariate associations show the correlation between two variables and indicate the direction and strength of the relationship between these two variables (e.g. between Factor 2 and L2 proficiency). Regression analysis, on the other hand, implies that an outcome depends on several variables that influence each other mutually. Regression aims to predict the outcome by considering these mutual influences. It is therefore possible for bivariate associations and the regression model to yield different tendencies.

Similar to our findings, other studies have documented weak or non-significant correlations between extrinsic motivation and L2 outcomes (Husfeldt & Bader Lehmann 2009: 18, Kreis et al. 2014: 31). Also, negative effects for extrinsic motivation on L1 literacy have been reported, e.g., by Wang & Guthrie (2004) for reading comprehension, Becker et al. (2010) for reading, or by Pajares et al. (2009) for writing. These results can be interpreted with reference to the SDT framework (Deci & Ryan 2002) which describes self-determination as a major driving force for achievement (see also Chapter 1, §4). Since extrinsic motivation is not usually related to self-determination, but rather stems from some external source, it is plausible that it impacts less on learning outcomes.

Several tentative conclusions can be drawn from our findings. A common factor for language aptitude and general learning abilities that is positively associated with L2 achievement implies that children with good results in cognitive tests are also good foreign language learners. Successful learners are therefore

likely to be “all-rounders”, i.e. children who do well in foreign language classes are probably successful students in general. Pimsleur (1966) may already have implied this as the PLAB included grades from subjects other than languages as an assessment criterion. Similarly, the data indicate that learners are unlikely to display difficulties related specifically and exclusively to foreign language instruction.

The effect of L2 Academic Emotion (anxiety, self-concepts and intrinsic motivation) on L2 outcomes is worth noting in relation to classroom practice. It highlights the importance of creating positive encounters with the target language that foster students’ enjoyment of L2 learning. This suggestion implies that ID variables that affect L2 achievement are dynamic and can be influenced by (adequate) learning conditions. While L2 motivation has been described as quite receptive to change (see also Chapter 8), the malleability of language aptitude, for instance, remains contested. Since components of language aptitude have emerged as vital players in children’s developing L2 competence in this chapter, as well as the predictive models (Chapter 4) and the variables underlying L1 German and L2 English development (Chapter 9), it would be worth knowing if they can be fostered to enhance foreign language learning. The stability of grammatical sensitivity and inductive ability has been investigated in the LAPS project and results are presented in Chapter 10.

5 Conclusions

We investigated the underlying structure of a broad set of ID variables involved in early instructed L2 learning. To this aim, a test battery of general learning abilities, language aptitude, and affective variables was administered to two different samples of primary school children aged 10–12 years. The research questions can be answered as follows:

Question 1: What is the underlying structure of the ID variables?

EFA yielded a 3-factor structure: Factor 1 subsumes general learning abilities and language aptitude and is named Cognition/Aptitude Factor. Factor 2 comprises extrinsic motivation, encouragement from adults and dedication. It is referred to as Extrinsic Motivation. Factor 3 involves emotions linked to L2 learning (intrinsic motivation, anxiety, and L2 self-concepts) and is therefore labelled L2 Academic Emotion. The model explains 40% of the variance, with Factor 1 explaining 16%, and Factors 2 and 3 contributing 12% each.

Question 2: Is the structure replicable in an independent sample?

The factor solution from the EFA was validated by means of CFA in an independent sample. This points to a robust factor structure, regardless of the population and target language. A distinction between language aptitude and general learning abilities was not supported by our data. Therefore, our results do not warrant for a special “language talent” that is independent of general cognitive abilities, nor do they indicate that some young learners are likely to struggle exclusively with foreign language instruction.

Question 3: What are the effects of the ID variables on L2 proficiency?

A regression analysis with the 3 factors as predictors and L2 proficiency as the dependent variable shows that L2 achievement is positively associated with Cognition/Aptitude and L2 Academic Emotion. When these two factors are controlled for, the extrinsic factor has a negative effect on L2 achievement.

Since findings from factor analysis cannot readily be generalized, we took care to provide conditions that maximize the informative value of our findings. Thus, we adopted a data-driven approach that included an exploratory and confirmatory analysis. Also, we selected two independent samples with – compared to other studies in the field of SLA – relatively large cohorts to optimize the validity of the factor structure. While the constructs assessed in the test battery remained unchanged, two tests pertaining to intelligence and phonetic coding ability were modified in accordance with advice from an expert panel. Also, L2 proficiency was operationalized differently in the two samples: as general language proficiency measured by C-Tests in LAPS I and as reading and listening comprehension in LAPS II. Results from the regression analysis were similar in both samples, suggesting a similar pattern of relationships between the three factors and different aspects of language proficiency. In the future, it would be interesting to extend to other areas, such as oral or written production.

Understanding L2 learning for the age group investigated here is relevant given the importance attached to early foreign language instruction in recent years. Therefore, we would encourage replications of the analyses presented in this chapter with samples from different populations learning different foreign languages.

References

- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: 10.18637/jss.v067.i01.
- Becker, Michael, Nele McElvany & Marthe Kortenbruck. 2010. Intrinsic and extrinsic reading motivation as predictors of reading literacy: A longitudinal study. *Journal of Educational Psychology* 102(4). 773–785.
- Bortz, Jürgen & Christof Schuster. 2010. *Statistik für Human- und Sozialwissenschaftler*. Berlin: Springer-Verlag.
- Brown, Timothy A. 2006. *Confirmatory factor analysis for applied research*. New York, London: The Guilford Press.
- Carroll, John B. & Stanley M. Sapon. 1976. *Modern language aptitude test – Elementary*.
- Csizér, Kata & Judit Kormos. 2009. Modelling the role of inter-cultural contact in the motivation of learning English as a foreign language. *Applied Linguistics* 30(2). 166–185. DOI: 10.1093/applin/amm025.
- Deci, Edward L. & Richard M. Ryan. 1985. *Intrinsic motivation and self-determination in human behaviour*. New York: Plenum.
- Deci, Edward L. & Richard M. Ryan (eds.). 2002. *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.
- Dörnyei, Zoltán. 2010. *Questionnaires in second language research*. 2nd edn. New York: Routledge.
- Grañena, Gisela. 2012. *Age differences and cognitive aptitudes for implicit and explicit learning in ultimate second language attainment*. College Park, MA: University of Maryland. (Doctoral Dissertation).
- Grañena, Gisela. 2013. Cognitive aptitudes for second language learning and the LLAMA language aptitude test. In Gisela Grañena & Michael Long (eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (Language Learning and Language Teaching 35), 105–130. Amsterdam: John Benjamins.
- Hallquist, Michael. 2018. *Introductory SEM using lavaan*. 2018 PSU Psychology R Bootcamp. https://psu-psychology.github.io/r-bootcamp-2018/talks/lavaan_tutorial.html.
- Heinzmann, Sybille. 2013. *Young language learners' motivation and attitudes: Longitudinal, comparative and explanatory perspectives*. London: Bloomsbury.
- Horwitz, Elaine, Michael B. Horwitz & Joann Cope. 1986. Foreign language classroom anxiety. *The Modern Language Journal* 70(2). 125–132.
- Husfeldt, Vera & Ursula Bader Lehmann. 2009. *Englisch an der Primarschule: Lernstandserhebung im Kanton Aargau*. Aarau: BKS.

- Kline, Rex B. 2011. *Principles and practice of structural equation modeling*. 3rd edn. New York: The Guilford Press.
- Kormos, Judit & Kata Csizér. 2008. Age-related differences in the motivation of learning English as a foreign language: Attitudes, selves, and motivated learning behavior. *Language Learning* 58(2). 327–355. DOI: 10.1111/j.1467-9922.2008.00443.x.
- Kreis, Annelies, Michaela Williner & Christoph Maeder. 2014. *Englischunterricht in der Primarschule des Kantons Thurgau: Schlussbericht der Evaluation*. Kreuzlingen: Pädagogische Hochschule Thurgau.
- Lenhard, Wolfgang & Wolfgang Schneider. 2006. *ELFE 1-6 Ein Leseverständnistest für Erst- bis Sechstklässler*. Göttingen: Hogrefe.
- Li, Shaofeng. 2016. The construct validity of language aptitude: A meta-analysis. *Studies in Second Language Acquisition* 38(4). 801–842. DOI: 10.1017/S027226311500042X.
- Liu, Meihua & Wenhong Huang. 2011. An exploration of foreign language anxiety and English learning motivation. *Education Research International* 2011. 1–8.
- Meara, Paul M., James Milton & Nuria Lorenzo-Dus. 2001. *Llama language aptitude tests: The manual*. Swansea.
- Noels, Kimberly A., Luc G. Pelletier, Richard Clément & Robert J. Vallerand. 2000. Why are you learning a second language? Motivational orientations and self-determination theory. *Language Learning* 50(1). 57–85.
- Nowicki, Stephen & Bonnie R. Strickland. 1973. A locus of control scale for children. *Journal of Consulting and Clinical Psychology* 40(1). 148–154.
- Oxford English Testing. 2013. *Oxford young learners placement test*. <https://www.oxfordenglishtesting.com/> (5 May, 2017).
- Pajares, Frank, Gio Valiante & Yuk Fai Cheong. 2009. Writing self-efficacy and its relation to gender, writing motivation, and writing competence: A developmental perspective. In Suzanne Hidi & Pietro Boscolo (eds.), *Writing and motivation*, 141–162. Amsterdam: Elsevier.
- Pekrun, Reinhard, Thomas Goetz, Wolfram Titz & Raymond P. Perry. 2002. Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist* 37(2). 91–106.
- Peyer, Elisabeth, Mirjam Andexlinger, Karolina Kofler & Peter Lenz. 2016. *Projekt Fremdsprachenevaluation BKZ: Schlussbericht zu den Sprachkompetenztests*. Tech. rep. Université de Fribourg. <http://doc.rero.ch/record/259287/> (9 February, 2018).
- Pimsleur, Paul. 1966. *Pimsleur language aptitude battery (form S.)* San Diego: Harcourt Brace, & Co.

- Rosseel, Yves. 2012. Lavaan: An R package for structural equation modeling. *Journal of Statistical Software* 48(2). 1–36.
- Ryan, Richard M. & Edward L. Deci. 2000. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology* 25. 54–67.
- Sasaki, Miyuki. 1996. *Second language proficiency, foreign language aptitude, and intelligence: Quantitative and qualitative analyses*. New York: Peter Lang.
- Skehan, Peter. 2019. Language aptitude implicates language and cognitive skills. In Peter Skehan, Adriana Biedroń, Richard L. Sparks, Zhisheng (Edward) Wen & Shaofeng Li (eds.), *Language aptitude advancing theory, testing, research and practice*, 53–75. New York: Routledge.
- Sparks, Richard L., John Patton, Leonore Ganschow & Nancy Humbach. 2011. Sub-components of second-language aptitude and second-language proficiency. *The Modern Language Journal* 95. 253–273.
- Stöckli, Georg. 2004. *Motivation im Fremdsprachenunterricht: Eine theoriegeleitete empirische Untersuchung in 5. und 6. Primarschulklassen mit Unterricht in Englisch und Französisch* (Pädagogik bei Sauerländer). Oberentfelden: Sauerländer Verlag.
- Vafaei, Payman, Yuichi Suzuki & Ilin Kachisnke. 2017. Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition* 39(1). 59–95. DOI: 10.1017/S0272263115000455.
- van Prooijen, Jan-Willem & Willem A. van der Kloot. 2001. Confirmatory analysis of exploratively obtained factor structures. *Educational and Psychological Measurement* 61(5). 777–792.
- Wang, Judy Huei-Yu & John T. Guthrie. 2004. Modeling the effects of intrinsic motivation, extrinsic motivation, amount of reading, and past reading achievement on text comprehension between U.S. and Chinese students. *Reading Research Quarterly* 39(2). 162–186.
- Weiβ, Rudolf H. 2006. *Manual Grundintelligenztest Skala 2 – Revision (CFT 20-R)*. Göttingen: Hogrefe.
- Wesche, Marjorie, Henry Edwards & Winston Wells. 1982. Foreign language aptitude and intelligence. *Applied Psycholinguistics* 3(2). 127–140. DOI: 10.1017/S0142716400006664.
- William, Revelle. 2018. *Psych: Procedures for personality and psychological research*. <https://www.scholars.northwestern.edu/en/publications/psych-procedures-for-personality-and-psychological-research>.
- Witkin, Herman A., Philip K. Oltman, Evelyn Raskin, Stephen A. Karp & Jack Demick. 2014. *Group embedded figures test manual*. Menlo Park: Mind Garden.

Chapter 4

Predicting L2 achievement: Results from a test battery measuring language aptitude, general learning ability, and affective factors

Raphael Berthele^a, Jan Vanhove^a, Carina Steiner^b, Isabelle Udry^{a,c} & Hansjakob Schneider^c

^aUniversity of Fribourg, Institut de Plurilinguisme ^bUniversity of Bern, Center for the Study of Language and Society ^cZurich University of Teacher Education

This chapter addresses the following questions: How, and how accurately, can the children's performance on a L2 English proficiency test administered around age 12 (T3) be predicted on the basis of test and questionnaire data that were collected one-and-a-half years earlier (T1)? The analyses suggest that a fairly simple equation can predict the score at T3 on new data with an average error of 1.8 points on a 20-point scale. This equation contains seven input variables available at T1. This accuracy corresponds to a R^2 of 0.58. The seven input variables comprise questionnaire-based affective variables, as well as aptitude measures and German (the pupils' school language) reading test data. Most of the heavy lifting, however, is done by an English proficiency test administered at T1, which by itself can predict new T3 data with an average error of 2.0 points ($R^2 = 0.41$).

1 Prognostic perspective on aptitude

Language aptitude testing originated in an interest in prognostic testing. As discussed in Chapter 1, early tests were designed to identify the strong learners within groups of students. This is also one of the research questions of our LAPS project: What is the prognostic value of the aptitude tests for the development of



Raphael Berthele, Jan Vanhove, Carina Steiner, Isabelle Udry & Hansjakob Schneider. 2021. Predicting L2 achievement: Results from a test battery measuring language aptitude, general learning ability, and affective factors. In Raphael Berthele & Isabelle Udry (eds.), *Individual differences in early instructed language learning: The role of language aptitude, cognition, and motivation*, 91–103. Berlin: Language Science Press.

DOI: 10.5281/zenodo.5464751 

proficiency in the foreign language? The second main question, regarding the underlying structure of individual dispositions to language learning, is discussed in Chapter 3. Whereas the investigation of dimensionality draws on and contributes to theory building and development regarding the construct of language aptitude, we opted to approach the question of prognostic modelling without a priori theorizing the relative weights of the constructs included in our investigation: We ask how the information gathered at T1 predicts English proficiency 1.5 years later at T3.¹ The overall aim is to extract a set of predictors that, taken together, would allow teachers to estimate the potential development of their students.

2 Modelling strategy

In a first step, our aim is to determine the model that most accurately predicts children's English proficiency at T3 when all variables assessed in the project are considered (we refer to this as the "no costs spared model"). In a second step, we attempt to find a model that is suitable for application in a classroom with comparable predictive value to the "no costs spared model". We refer to such a model as the "cheap model". We require that such a model be based on tasks that can be conveniently administered within a 45 minute lesson and evaluated easily by a non-specialist teacher. To this aim, we compared the comprehensive (i.e. no costs spared) model against simpler models, which included background information readily available to the teacher and short tests from our test battery.

The main steps of the process will be summarized in the following. For full details on how we went about building the predictive models, the reader is referred to the technical report (Vanhove 2021).

1. Split up the dataset into a training and a test set. The training set was used for trying out different models and for gauging the strength of these different models. Based on the results, the optimal, final model was selected. The test set was used for validating the selected model.
2. Compute scores for constructs such as intrinsic motivation and locus of control. Because of their conceptual and statistical ease of use, we preferred equally-weighted scales wherever they seemed reasonable.

¹It is, of course, also possible to fit models that predict a student's performance on the English test at the second data collection using information available at the first or that predict their performance at the third data collection using information available at the first and second. Indeed, in preliminary analyses we also fitted such models. We limit our discussion here to the models we consider the most relevant ones.

3. Exclude students that are not of interest for the present research question. These are students whose native language is English or who were exempted from English classes. Pupils who did not take the T3 English test were also excluded from the analyses.
4. Reduce the number of predictors. We removed predictor variables with little variance in the training set and we removed one predictor (ELFE total) showing a very strong intercorrelation (0.94) with ELFE word in the training set.
5. Impute missing data. Missing values in the predictor variables that were retained were imputed using the median of the available values of the same variable.
6. Fit models and cross-validate them in the training set. In order to gauge these models' predictive strength without turning to the test set, cross-validation was applied (§3.2). This is a technique that essentially mimics the partitioning of the overall dataset into a training and test set (§3.1). We fitted a whole family of models:
 - a. A “no costs spared” model with all variables assessed.
 - b. Two simple baseline models so that we could get a sense of how much better the “no costs spared” model actually performed in cross-validation.
 - c. Four “cheap” models that could potentially be applied in classroom settings.
7. Select the final models. The final “no costs spared” and the final “cheap” models were decided on by the research team based on the candidate models' likely predictive strength (estimated by cross-validation) and, in the case of the “cheap” models, the costs involved in obtaining the predictor information required.
8. Assess the predictive strength of the final model using the test set. The final model was refitted on all of the training data and its predictive strength was then tested on the test set. Importantly, the model's parameters and settings were not re-estimated or tweaked using the test set.

3 Data partition and cross-validation

3.1 Training and test sets

Some analyses in this project are exploratory by nature (e.g., the exploratory factor analysis in Chapter 3). Exploratory analyses entail the substantial risk that the models tightly fit the dataset analyzed but do not generalize well beyond it. To offset this risk, we partitioned the dataset analyzed in this chapter into a training set and a test set (see Kuhn & Johnson 2013).

The training set was used to conduct all exploratory analyses and to decide on such matters as data transformations, the calculation of construct scores, missing data imputation, model specification – in a nutshell, any step in the analysis that requires the analyst to take a decision. Once a suitable predictive model was agreed upon, its predictive power was tested on the test set. Crucially, the chosen predictive model was not re-estimated using the test set data.

To respect the hierarchical nature of the data (children in classes), the test and training sets were not random subsets of the children in the study, but rather (largely) random subsets of the classes in the study (see Roberts et al. 2017). In this way, we could account for the clustering of the pupils in classes when estimating the prediction error in our models. Specifically, from the 17 grade 4 classes at T1, 5 were selected to comprise the test set: the smallest class (Class 4, with 5 grade 4 pupils at T1)² as well as four randomly picked classes. Similarly, from the 19 grade 5 classes at T1, 6 were selected to comprise the test set: the smallest class (also Class 4, with only 1 grade 5 pupil at T1) as well as five randomly picked classes. The remaining 12 grade 4 and 13 grade 5 classes comprised the training set.

3.2 Cross-validation

When trying out different models on the training data, we used cross-validation to estimate how well the models would work for new data.³ This was done to ensure that overzealous data exploration and model fine-tuning would not result in a model that fits the training data well but stands little chance of predicting the test data (see Kuhn & Johnson 2013, Yarkoni & Westfall 2017). In cross-validation, the training data is split up into a number (k) of folds, and models are fitted on $k-1$ folds and then used to predict the outcome in the remaining fold. This process is repeated k times, each time leaving out a different fold. The result is k estimates

²The sometimes small number of pupils per class is a consequence of some of the classes being mixed grade.

³This section is adapted from Vanhove et al. (2019).

Table 1: Training and test sets. *Note:* The number of classes sums to 36 rather than 32 because four classes had pupils from both 4th and 5th grade at T1. Only pupils for whom T3 English scores were available were included in the predictive models; for the final models, we only included participants who also had T1 English scores.

Set	Cohort	Classes	English data at	
			T3 only	both T1 and T3
Training	4th grade at T1	12	169	154
	5th grade at T1	13	187	177
Test	4th grade at T1	5	70	65
	5th grade at T1	6	85	80

of the models’ predictive accuracy on data not used for fitting the model that can then be averaged. The folds were not constructed randomly, since we need to account for the dependency structure in the data (pupils in classes). Therefore, we opted for block cross-validation, using each class as a separate fold.

Figure 1 (page 96) illustrates the principles behind the partitioning of the data and block cross-validation.

3.3 Metrics of model performance

The root mean squared error (RMSE) was used to adjudicate between different models. The RMSE can be interpreted as being roughly – but not quite – the average difference between a model’s predictions and the observed values. (In the same way that a standard deviation can be interpreted as being roughly – but not quite – the average difference between the observations and their mean.) The interpretation of the mean absolute error (MAE) is simpler: It is the average (mean) difference between a model’s prediction and the observed values. We report both metrics in this chapter.

Many readers will be more familiar with the R^2 metric of (so-called) “explained” variance. Several problems beset R^2 , but perhaps most important of all is that R^2 , as it is traditionally computed, does *not* estimate how well the model itself would capture the variance in a new sample. Instead, it estimates (at best) how well a *newly estimated* model would capture the variance in a new sample.⁴

⁴There exist different ways of computing R^2 (Kvalseth 1985). For ordinary regression models, these all yield the same result. However, when the model is used to predict observations that were not used when fitting the model, they do not. One popular method for computing R^2 is

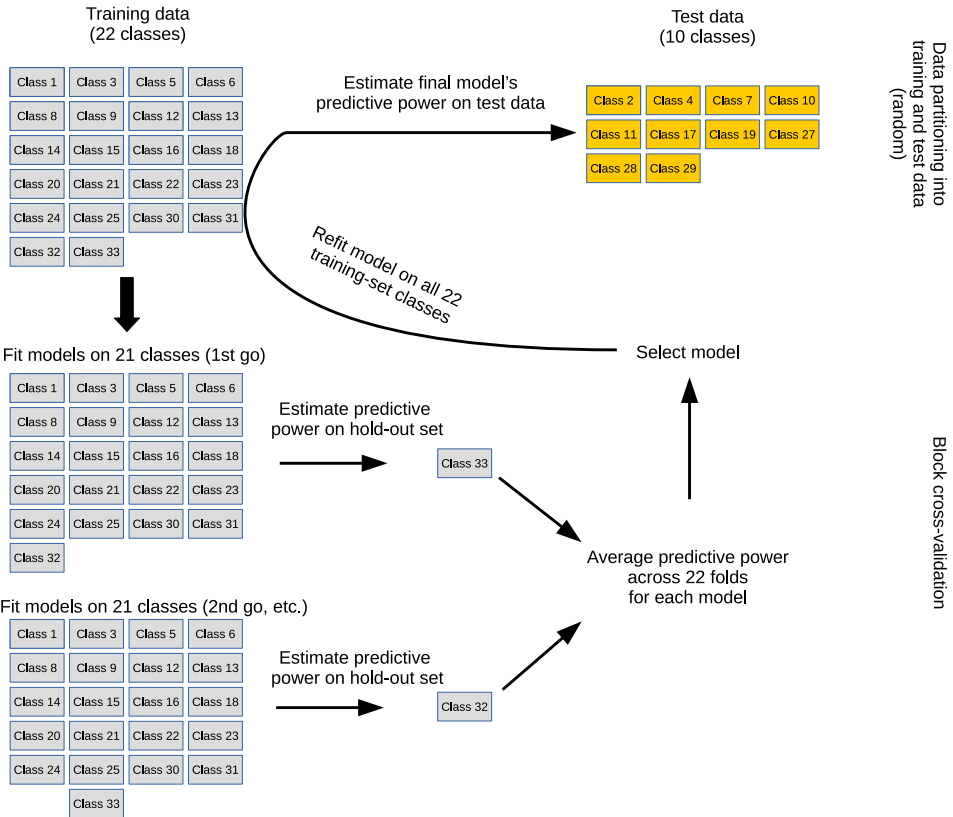


Figure 1: Illustration of how the data were partitioned into a training ($N = 356$) and a test set ($N = 155$) and of how block cross-validation works. Only two iterations of block cross-validation are shown; in reality, 22 took place for each model, each time leaving out a different class. Figure based on Figure 3 in Vanhove et al. (2019).

to square the correlation between the predicted and observed values. This is problematic since the correlation between predicted and observed values can be excellent even if the former corresponds poorly to the latter (e.g., the values 1, 2, 3 correlate perfectly with the values 2000, 4000, 6000 but correspond poorly to them). We therefore computed R^2 as the proportional decrease in the residual sum of squares relative to a baseline model without any predictors. Such a model predicts each new observation to be equal to the mean of the training data (footnote adapted from Vanhove et al. 2019).

4 Selection of the “no costs spared” model

As shown in Table 1, the training set for T3 comprised 169 4th graders and 187 5th graders. The test set comprised 70 4th graders and 85 5th graders.

To fit the “no costs spared” model, all available T1 information, from all possible sources, was allowed to enter into this model, without regard to how difficult or costly it was to collect this information.

To arrive at the final model in this category, a host of models were fitted on the training data. These included multiple linear regression, robust regression, ridge regression, elastic net, multivariate adaptive regression splines, generalized additive models, partial least squares regression, k-nearest neighbors, regression trees, random forests, support vector machines, stochastic gradient boosting, and Cubist.⁵ A multiple linear model with seven predictors and no interactions (Table 2) performed roughly on par with the more complex approaches in cross-validation. When fitting the final model, we only took into account participants who had T1 English test scores. The model’s estimated coefficients are shown in Table 2.

We want to draw the attention of any reader who wishes to use this model for understanding (as opposed to predicting) foreign-language learning to what Breiman (2001) calls the “Rashomon effect”: While the presented model worked best in cross-validation, a number of models with different predictors fared only slightly worse. Consequently, one would be jumping to conclusions if one said that the seven predictors listed in Table 2 are important in foreign-language learning and the others are not. The performance of the more complex models in cross-validation can be consulted in the online materials⁶.

Second, we fitted two simple baseline models so that we could get a sense of how much better the “no costs spared” model actually performed in cross-validation. The first baseline model was a “no predictor” model, which predicted each unseen data point to be equal to the mean of the seen data points. The second was an “English-only” model, which only contained the participants’ T1 English test score as the predictor of their T3 English test score. This was done because the English score at T1 unsurprisingly explains the largest share of variance of English at T3 since it taps into the same construct.

In cross-validation, the “no costs spared” model (Table 2) with seven predictors fitted the data better than the baseline models, the residual sum of squares is reduced by about 58% relative to an intercept-only model (i.e., $R^2_{\text{RSS}} = 0.58$, 95% CI: [0.49, 0.66]).

⁵For details on the architecture of these models, see www.osf.io

⁶<https://osf.io/ha7s2/>

Table 2: Multiple linear regression model for predicting T3 English scores. *Note:* Missing predictor data were imputed using median imputation using the full training set data. Median = the predictor’s median in the training set (used in imputation). Estimate = the estimated regression coefficient for the predictor. SE = the naïve standard deviation for the estimated regression coefficient; naïve meaning that its computation did not take into account the fact that this model was selected for its performance in cross-validation.

Term	Median	Estimate	SE
Intercept		0.045	1.4
English T1	49	0.093	0.010
Grade at T1	5	−0.55	0.31
Intrinsic motivation	3	0.48	0.22
Self-concept English	3	0.89	0.22
German (ELFE sentences/minute)	7.33	0.49	0.08
Grammatical sensitivity (gra)	15	0.047	0.023
Inductive ability (ind)	5	0.16	0.05

The “no costs spared” model’s root mean square error (RMSE) in cross-validation was 2.24 (95% CI: [2.02, 2.47]), and its mean absolute error (MAE) in cross-validation was 1.77 (95% CI: [1.60, 1.95]). For reference, an intercept-only model yielded a RMSE of 3.69 and a MAE of 2.93. A linear model with a single predictor, viz., the participants’ English score at T1, was also fitted and cross-validated. This model yielded $R^2_{\text{RSS}} = 0.42$, RMSE = 2.61 and MAE = 2.03. These results from the cross-validation analysis, also shown in Figure 2, suggest that there is a considerable gain in predictive performance when English at T1 is used, and some further gain when, in addition to English at T1, the other six predictors from Table 2 are included.

When applied to the test set, the linear model with seven predictors reduced the residual sum of squares by about 62% relative to the intercept-only model (i.e., $R^2 = 0.62$, 95% CI: [0.52, 0.70]). Its root mean square error (RMSE) was 2.32 (95% CI: [2.02, 2.60]) and its mean absolute error (MAE) 1.85 (95% CI: [1.63, 2.08]).

These results are summarized in Figure 2.

5 Selection of the “cheap” models

Next, we fitted four models that include sets of variables that are less costly to acquire or even completely “free” in the sense that the required information on

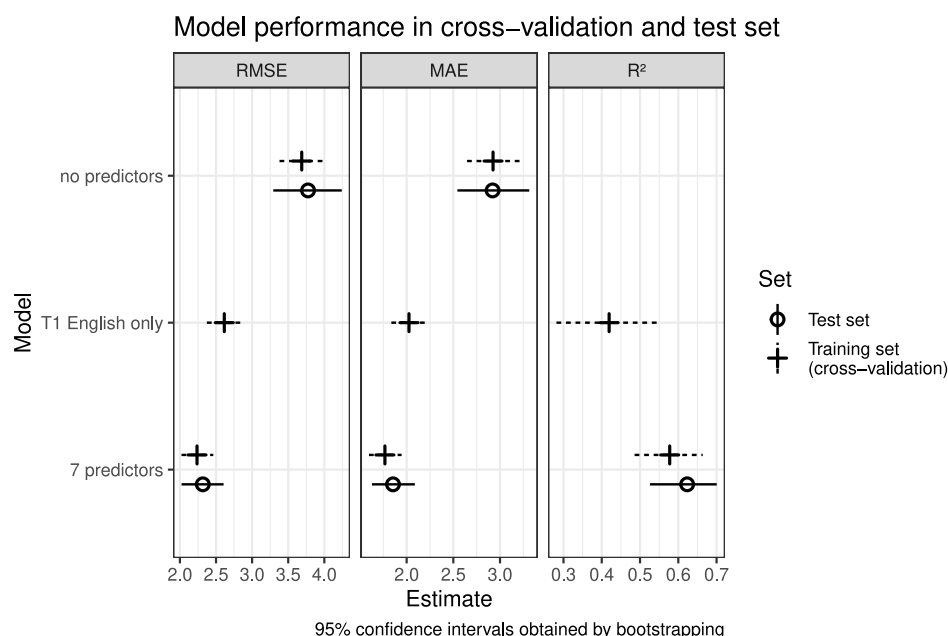


Figure 2: Performance of the chosen model relative to two baseline models. Note: The R^2 value for the intercept-only model is not shown as it is 0 by definition. The 95% confidence intervals were obtained by bootstrapping the 22 cross-validation estimates or by bootstrapping the observed and predicted test set values and recomputing the estimates (percentile approach). The English-only model wasn't applied to the test set.

the pupils is usually a given in the school context. We refer to these models as “cheap” models. They could be used in a classroom setting without having to collect data during four full lessons, the required time for the full LAPS test battery. The rationale here is to explore the possibility of predicting foreign language achievement based on information that is not complicated to get.

Free variables encode information that is available to teachers anyway as opposed to, say, a computer-based test of working memory. Information included as “free variables” were whether students had additional support in class, grade, and whether their L1 was German. This only involves information that wouldn't lead to discrimination based on sex (no gender variable) or possibly socioeconomic background (no SES variable from the parent questionnaire).

Relatively “cheap” variables are measures that are easy to take (paper and pencil instruments in the case of the aptitude tests, or questionnaire items on motivation). In this selection process, we also took into consideration a small set

of measures that are somewhat less straightforward to acquire (since they need to be purchased and/or adapted), such as language tests and aptitude tasks, but that are highly relevant to the central constructs of our investigation. Included here were thus T1 English score (OYLPT), T1 German score (ELFE), T1 motivation questionnaire-based construct scores, and the inductive ability score (PLAB form 4).

Four cheap models were fitted. The first one includes English at T1 and all the free variables mentioned above. The second includes reading skills in the school language (German) and all free variables. The third includes the motivational items regarding English and the free variables, and the last one motivational items and the adaptation of the PLAB subtest of inductive learning. The performance of these four models on the test set are given in Table 3.

Table 3: The four cheap models and their performance on the training and (for two of them) test sets. Tr: Training set; Te: Test set. *Note:* The training set estimates were obtained through cross-validation.

Model no. & included variables	RMSE		MAE		R ²	
	Tr	Te	Tr	Te	Tr	Te
1 English T1 + free variables	2.48	2.52	1.99	1.96	0.47	0.55
2 L1 German (ELFE) + free variables	2.69		2.19		0.39	
3 Motivation + free variables	2.83		2.27		0.34	
4 Motivation + ind + free variables	2.69	2.82	2.16	2.24	0.40	0.46

Our research group then discussed which ones of these cheap models should be validated on the test set. For the reasons spelled out above, most importantly to avoid over-fitting, we wanted to select a maximum of two models. Given the known strong association of the English at T1 test with our outcome variable at T3, the first cheap model was to be retained. The second model selected was the next best model according to the MAE and RMSE performance on the training set. This model includes motivational items and the inductive ability test based on a form of the PLAB.⁷ The first and fourth model were then applied to the test set. When applied to the test set, the cheap model 1 (English at T1 and all of the free variables) did better than the cheap model 4 (motivation, inductive ability and free variables). As a reminder, the free variables contain information that

⁷With the kind permission of Charles Stansfield, LLTF, we adapted this form of the PLAB without having to pay any fees. If, however, a German version of PLAB should be developed in the future, this would most likely not be free to use.

is often known anyway to teachers (and if not readily available), e.g. whether students had additional support in class, the pupils' current grade, and whether one of their L1 was German. This cheap model 1 has an R^2 of 0.47 on the training set and of 0.55 on the test set.

6 Discussion

The general goal of the analysis presented in this chapter was to assess the possibility to prognosticate foreign language development in primary school children. As discussed in Chapter 1, prognostication was what inspired the first practitioners and scholars to develop modern language aptitude tests: Predicting the success in learning a new language would help select the “apt” learners and prevent spending time and money on not so apt individuals. Most of this research and development focused on (young) adults and adolescents.

Our primary goal was not to provide an instrument for selection – foreign language education in the context investigated here is compulsory anyway. However, prognostication can serve other purposes as well (see Chapter 11 for a discussion).

Our attempt to investigate prognostication took several alternative paths. The first path, the “no costs spared” approach, took all information available at T1 into consideration. The multi-step modelling procedure explained in this chapter yields a model that we expect to predict the English score (on a scale from 0 to 20) of new data with a mean absolute error of about 1.8 points. Whether this is a good or a bad model performance depends on the criteria one wishes to apply. It certainly shows that the development in the foreign language is not simply random but is constrained by some of the constructs known and used in research on individual differences in language learning. Among the seven predictors in the model that showed the best performance on our test set, we find both emotional/motivational variables (intrinsic motivation, self-concept) and language-related variables (grammatical sensitivity [based on MLAT-E part 2] and inductive ability [based on PLAB, form 4], but also German reading proficiency).

The fact that the English skills tested one and a half years before T3 is the best predictor for the same skill is not surprising. But it also shows that, within the time interval covered by this study, relative differences in English skills are retained (see Chapter 10 for an analysis of the intra-individual stability across time of other constructs).

As discussed above, other models with different sets of predictors show almost the same performance (Rashomon effect). The list of variables of our best model

therefore is not to be read as a final list of predictors of success in foreign language learning, implying that the variables not retained are not important. What our analyses merely show is that these measures are robustly associated with foreign language development, even when we apply cautious modelling that aims at avoiding over-fitting the data. In support of the relevance regarding the variables just mentioned, we can also refer to the cross-sectional analysis discussed in Chapter 3 where we show that these variables are part of the two constructs that are positively associated with English skills at T1 (we labelled these factors *Cognition/Aptitude* and *L2 Academic Emotion*).

Among this first set of models that yield the optimal “no cost spared” model, we found that a simple model with T1 English as its sole predictor performs roughly on par with the other, more complex models (state estimate of about 2 on a scale 0–20).

After this first analysis, we took a less maximalist but more pragmatic perspective on prognostication. This involved fitting four “cheap” models to the data sets. In certain pedagogical contexts, it can be useful to assess individual differences in foreign language learning with relatively simple tests. Thus, we compared different sets of tests with respect to their prognostic values. All of them can be administered in a classroom setting and do not take up more than 45 minutes. These models include only variables that are already available to the teacher or that do not require complicated testing procedures. From these cheap models, the one with the T1 English test score plus the free variables and the one with motivation plus the inductive ability test from the PLAB plus the free variables were selected and applied to the test set. Based on a comparison of the performance of these models, it appears reasonable to use the Oxford Young Learners Placement Test plus the set of information we termed the “free variables” for prognostic use. This test of English skills is not freely available (it currently costs around £5 per pupil) and it would be a matter of choice on the part of the teachers or schools whether they would wish to consider such a test.

The analysis in this chapter aims to identify the best combination of measures taken at T1 to predict English skills at T3. The models discussed here, and their variable prognostic performance, are not designed with the goal to disentangle the different dimensions of predispositions for language learning. This was the main goal of the analyses in Chapter 3. Despite the different questions asked in Chapters 3 and 4, it seems reassuring that when looking at the list of variables in our “no costs spared” model (Table 2), we find tests that all load onto one of the two factors that are positively associated with English skills at T1, that is measures we subsumed under the two labels cognition/aptitude and L2 academic

emotion in Chapter 3. As in that preceding chapter's analyses, extrinsic motivational constructs are not first in line when their association with English skills is investigated.

References

- Breiman, Leo. 2001. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science* 16(3). 199–231. DOI: 10.1214/ss/1009213726.
- Kuhn, Max & Kjell Johnson. 2013. *Applied predictive modeling*. New York: Springer. DOI: 10.1007/978-1-4614-6849-3.
- Kvalseth, Tarald O. 1985. Cautionary note about R^2 . *The American Statistician* 39(4). 279–285. DOI: 10.2307/2683704.
- Roberts, David R., Volker Bahn, Simone Ciuti, Mark S. Boyce, Jane Elith, Gurutzeta Guillera-Arroita, Severin Hauenstein, José J. Lahoz-Monfort, Boris Schröder, Wilfried Thuiller, David I. Warton, Brendan A. Wintle, Florian Hartig & Carsten F. Dormann. 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* 40(8). 913–929. DOI: 10.1111/ecog.02881.
- Vanhove, Jan. 2021. *Language aptitude in primary school*. Tech. rep. Fribourg, Switzerland. <https://osf.io/d9gnh/>.
- Vanhove, Jan, Audrey Bonvin, Raphael Berthele & Amelia Lambelet. 2019. The lexical richness of short French, German and Portuguese texts written by children. DOI: 10.17605/OSF.IO/VW4PC.
- Yarkoni, Tal & Jacob Westfall. 2017. Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science* 12(6). 1100–1122. DOI: 10.1177/1745691617693393.

Chapter 5

Dispositions for language learning and social differences

Raphael Berthele


University of Fribourg, Institut de Plurilinguisme

The acquisition and the use of languages are cognitively and socially shaped. Specific groups (migrants, socially lower strata) have been shown to perform consistently worse in (language) learning in western educational systems. This chapter investigates the association of social variables with language skills and with the dispositions for language learning discussed in this book. The chapter briefly reviews current assumptions and findings on the role of material, educational and cultural characteristics of pupils' family backgrounds for tutored language learning in school contexts. Using structural equation modelling techniques, several models are compared to assess the role of specific background characteristics. The comparisons involve the relative impact of economic vs. cultural predispositions, and the impact of migrant status and family languages that are different from the school language. The analyses show that it is empirically adequate to distinguish between cultural/educational and material family background. Moreover, once these two dimensions are part of the model, being a pupil with a migrant background or being multilingual does not make any difference with respect to English skills.

1 Why the Social matters

The goal of this chapter is to supplement the insights into the aptitude construct and individual difference (ID) variables discussed earlier in this book with variables that operationalize social information regarding the pupils' families. The questions answered in this chapter pertain to the role of family-related variables in the learning of a foreign language (see Chapter 2, Table 2. for more details about these variables). What roles do parents' education and income play in the



Raphael Berthele. 2021. Dispositions for language learning and social differences. In Raphael Berthele & Isabelle Udry (eds.), *Individual differences in early instructed language learning: The role of language aptitude, cognition, and motivation*, 105–124. Berlin: Language Science Press. DOI: 10.5281/zenodo.5464753 

learning of English skills? To what extent do the pupils' families' economic and cultural predispositions affect English skills and the factors explaining and predicting English skills? How much does it matter whether or not the children speak the language that is the medium of instruction (German) at home? Are speakers of other first languages overwhelmed by learning the local language as well as additional foreign languages? And lastly: How much does it matter whether the pupil or her parents are born in Switzerland or not?

Language (first, second, foreign) learning and use are social practices that rely on the human language making capacity (Slobin 1985). In the previous chapters, we addressed several different cognitive and affective aspects that have an impact on language learning. Although a factor representing SES (socioeconomic status) was included in the “no costs spared model”,¹ we did not further investigate social or other background information in our explanatory and prognostic analyses. This does by no means imply that we consider the social factors that shape (language) learning irrelevant. Social differences, for example regarding the pupils' families' levels of income or education, must not be excluded from the inquiry of learner differences. A sociological view on education, as discussed in this chapter, is an important counterweight to the psychometric approach that bears the risk of overemphasizing the individual while neglecting the structures in which the individuals do or do not unfold their potential. Individual differences in language learning are affected by the interplay of background and the educational system.

Although the influence of family background variables and the way they interact with characteristics of the educational systems are not at the core of the LAPS project, it would be naïve to pretend that the inquiry into IDs or language learning ability can be done in a sociological vacuum. Therefore, in this chapter I will address the questions raised above drawing on the larger sample of our project (LAPS II) using the technique of structural equation modelling.

1.1 Starting (too) simple: Money and English

We have shown in Chapter 4 that the English test scores at T1 are the strongest predictor for English skills at T3. Since English at T1 seems so important, to what extent are these skills associated with family background variables? In a preliminary attempt to investigate this association, we can plot the bivariate relationship of the global English test score at T1 and a variable that operationalizes the parents' income (see Table 1 for details). The left panel in Figure 1 shows this association.

¹See p. 37 in the technical report (Vanhove 2021). The factor did not turn out to be prognostically explanatory once other predictors were accounted for.

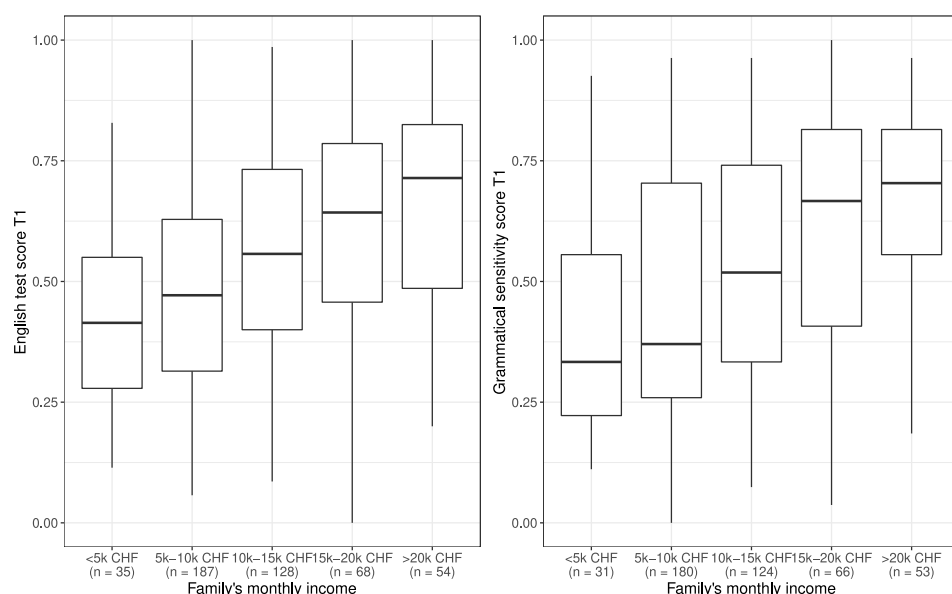


Figure 1: Associations between earnings and a) English test score and b) Grammatical sensitivity score at T1. There are 66 missing answers for the income question in the left and 64 missing answers in the right panel.

The positive association of parental income and the development of language abilities comes as no surprise. The positive association between the socioeconomic status of a child's family and their school performance has been widely investigated in the field of sociology of education (see Entwisle & Alexander 1992 for a classic study on the topic, Fernald et al. 2013 for a recent example on very early differentiation, Häberlin et al. 2005 for an investigation of the Swiss situation).

Figures 1 a and b show monotonic associations between the parents' income on the one hand and general proficiency in English at T1 and with language analysis as measured by the *words in sentences* task of our adapted subtest from MLAT-E on the other hand. While the English score represents the main outcome variable of the researcher interested in foreign language learning, the MLAT-E subtest represents a classic, prototypical language aptitude measure. The Figures show that not only the skills themselves but also abilities underlying the development of such skills in the target language co-vary with income.

A second association – and one that is often embraced in public debates on the multilingual curriculum – concerns teaching foreign languages to children with an immigrant background who do not speak the local language in their families.

Critics of the current multilingual curriculum that includes two compulsory foreign languages in primary school argue that in particular these second language learners are overwhelmed by the task of learning the local language and medium of instruction as well as two foreign languages such as English or French (Kübler et al. 2014). Figure 2 shows the association of English scores with groups of children: (a) children who do or do not speak German as the only or one of their languages at home and (b) children who were born in Switzerland or abroad.

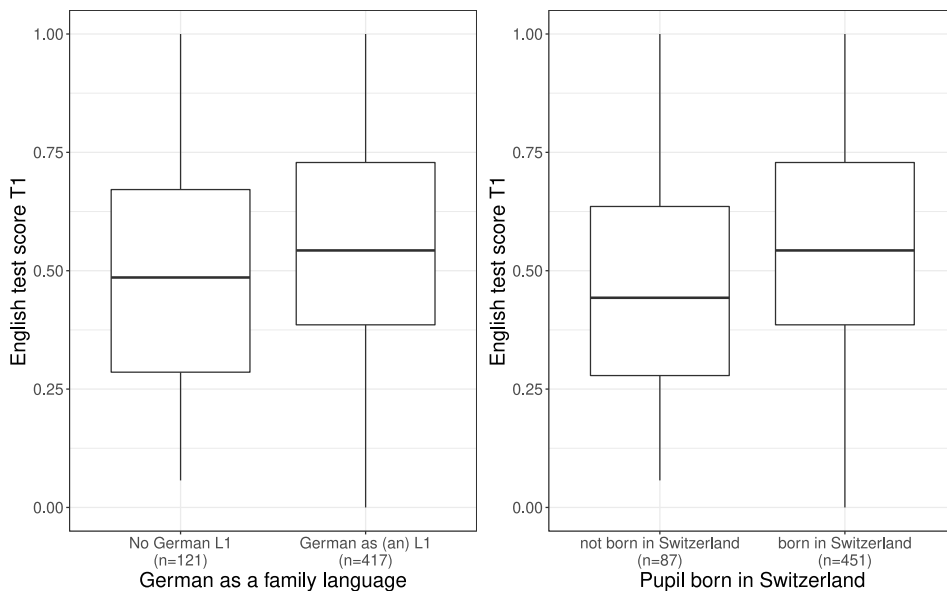


Figure 2: English scores at T1 and either German as a home language (left) or country of birth of the children (right).

As expected, the central tendency for the German speakers and for the Swiss born children is higher, which seems to confirm the expectations expressed above.

In this chapter, I add information such as the four variables plotted in Figures 1 and 2, that is, information on the family background of the pupils in the LAPS-II data, to the analysis of language aptitude, cognitive and affective learner dispositions. In doing so, I use the operationalizations and the statistical modelling presented in Chapter 3. As in these previous analyses, children who speak English at home and who have spent substantial parts of their lives in an English-speaking context were excluded. In addition to the factors identified in Chapter 3, I add two constructs related to the individual background of each learner. These constructs are arguably systematically associated with language learning and

cognitive abilities in general, as laid out in the next section. A more complicated structural equation model than the one discussed in Chapter 3 will therefore be fitted to the data of the first measurement point. The goal is to shed light on the associations between these additional background variables and the language-related constructs that are at the core of our inquiry.

1.2 Social background and educational success

This chapter focuses on explaining variance in English skills at the first measurement time in the LAPS-II project. In Chapter 4, we asked the question which factors allow us to predict skills in the target language at T3. When controlling for many potentially relevant factors, we found that English at T1 turns out to be the most informative predictor variable for these skills. Adding social information, however, does not increase the prognostic accuracy.

Sociolinguists and sociologists of education have accumulated a great wealth of evidence on the systematic associations of social class and language learning and use (see Avineri & Johnson 2015 for an overview). Most of the evidence concerns first or second language learning, studies of the social conditioning of foreign language learning are relatively scarce (but see DESI-Konsortium 2008 for a study that includes social information). If the influences on foreign language learning are similar to those observed in school learning in general, similar associations of the social background characteristics with the outcome variable are to be expected.

The learning of two foreign languages in our primary school context is not an elite phenomenon. All children in state run schools and in private schools complying with the national curricula start learning two foreign languages during the primary school years. Our sample thus is not a self-selected sample of pupils following a special curriculum.

As shown in the literature on inequalities in education, groups can be discriminated against based on their gender, social class, ethnicity or race (Lucas & Irwin 2018). In our context, two particularly important categories are systematically confounded: Pupils from immigrant families speaking other languages than German at home tend to be pupils from low-income families, too (Kronig 2003). As a consequence, pupils with a migrant background cumulate risk factors related to class and to language. Investigating additional language learning in these children means testing competing hypotheses from different fields: From the sociological point of view, immigrant children are expected to perform worse in school in general for the reasons spelled out above. From the point of view of third language acquisition theory, with the exception of pupils who start simul-

taneously learning both German as an L2 and English as L3,² such learners are expected to have fewer problems learning an additional language since they have the necessary learning apparatus already up and running due to their previous second language learning experiences (Herdina & Jessner 2002).

Given these contradictory expectations from different disciplines, it is useful to test the impact of social status and L2 learner status in a comprehensive model.

1.3 Educational, cultural and material indicators as background variables

The background information used in this analysis consists of variables gathered via a parental questionnaire (see Table 5 in Chapter 2 and Chapter 6 in Vanhove 2021 for more details). They encompass information on the origins and first languages of the parents as well as on economic and educational aspects of the families.

Socioeconomic status arguably should not simply be expressed via economic indicators such as the earnings used in Figure 1. An influential approach to status and its reproduction is the theory of different capitals proposed by Bourdieu (1979). Terms such as cultural and economic capital (among other types of capital) are used in different, evolving ways by Bourdieu himself and by other scholars, and it is impossible to give a detailed account of the theory of the different capitals here (see Farkas 2018 for a discussion citing the most recent literature). Bourdieu's theorizing on the different instantiations of capital is rather complex and also controversially debated in sociology (see Riley 2017 for an example). I do not claim that mapping a handful of manifest variables to two constructs labelled *economic* and *cultural capital* does justice to these debates. The insight, however, that social status and legitimacy is not only a matter of money certainly echoes parts of Bourdieu's early thinking about dimensions, mechanisms, and resources mobilized in the reproduction of social inequalities. Analogous distinctions between material dispositions of study participants on the one hand and cultural and educational dispositions on the other are often made by scholars representing different disciplinary perspectives and thus, they seem uncontroversial. On these grounds, it seems reasonable that the distinction should also inform the construction of the structural equation model in this chapter. In the remainder of this chapter, therefore, I will use the terms economic and cultural predispositions (*econ_p* and *cult_p*) to refer to two dimensions of the pupils' family background.

²Within the group of pupils who do not speak German as a family language, 3 out of 72 4th graders and 2 out of the 49 5th graders did not go to a German-language school in the Canton of Zürich the year before.

Table 1 lists the new variables used in the following analysis. The manifest variables operationalizing the constructs academic emotion (*acad_emo*) and *cognition* are discussed in Chapter 3.

For further statistical treatment, all numeric variables were z-transformed, i.e., they were centered at their respective sample mean and then divided by their respective sample standard deviation. The participants were selected following these criteria (see Chapter 14 in Vanhove 2021 for details): no native speakers of English, no substantial exposure to English beyond school input, English T1 data available. This left a dataset of 538 pupils.

Based on these variables, it was possible to assess the impact of economic, cultural, and linguistic characteristics of the pupils' families on the constructs at the core of the LAPS project and to answer the questions raised at the beginning of this chapter.

2 Modelling strategy

The data were modelled using the structural equation modelling function (*sem()*) of the *lavaan* package (Rosseel 2012) for R. Supplementary materials with the code and details of the analyses can be downloaded from osf.io.³ The latent constructs are identical to the first and third factor in the confirmatory factor analysis described in Chapter 3 (cf. Figure 1 in Chapter 3; see supplementary material to Chapter 3, modelling variant 4, for details).

English: English proficiency at T1 as measured by the OLYPT test. Two manifest scores are available, language use (“use”) and listening comprehension (“listening”).

Academic emotion (*acad_emo*): The same factor as the one described in Chapter 3; manifest measures of English self-concept (“sel”), foreign language anxiety in the English classroom (“anx”), and intrinsic motivation to learn English (“int”).

Cognition: The same factor as the one described in Chapter 3; manifest measures of intelligence (“iq”), visual and verbal working memory (“vis”, “vem”), Phonetic coding ability (“pho”), grammatical sensitivity (“gra”), inductive ability (“ind”), field independence (“fie”), and German reading (“ger”).

To this list I add two constructs related to the individual background of each learner (see variables listed in Table 1):

³https://osf.io/kdxc7/?view_only=d6b0409de06f4d5cb82f2678471af56b

Table 1: The background variables elicited with parents' questionnaires that are used in this chapter

Node label	Variable	Comment
Swiss	Was the pupil born in Switzerland?	Switzerland the pupil's country of birth (not citizenship)
educ_f	Highest level of father's education	Primary school, Secondary school, Apprenticeship, "Berufsmatura" (a higher vocational school leaving certificate), Academic Maturity, Higher Vocational Education or University Degree
educ_m	Highest level of mother's education	(1) <5k, (2) 5–10k, (3) 10–15k, (4) 15–20k, (5) >20k
earnings	Family's monthly income in CHF	1 (not at all sufficient), 2 (rarely sufficient), 3 (sometimes sufficient), 4 (usually sufficient), 5 (always sufficient)
bills	"Is there enough money to pay the monthly bills?"	
savings	"Is there enough money for saving?"	
holidays	"Is there enough money for travelling and holidays?"	
medical	"Is there enough money for medical and dental care?"	
books	"How many books are there in your household?"	(1) 0–10, (2) 11–25, (3) 26–100, (4) 101–200, (5) 201–500, (6) 500+
L1_Ger	German (a) family language	Yes if German (a) family language
ParentsSwissBoth	Country in which pupil's parents were born	Yes if both parents born in Switzerland
ParentSwiss		Yes if at least one parent born in Switzerland

Economic predispositions (econ_p): Manifest measures are “savings”, “bills”, “medical”, “holidays”, and “earnings”.

Cultural predispositions (cult_p): Manifest measures are “L1_Ger”, “Swiss”, “books”, “educ_f”, and “educ_m”.

2.1 Assumed causalities and correlations

We expect the factor *academic emotion* to be correlated with family background. Homes in which there is greater interest in culture, literature, and science arguably tend to foster more positive emotional conditioning towards school and learning in general, and most likely also towards learning English as a powerful vector of culture and science. We do not, however, assume a unidirectional causality from the background variables to the construct *academic emotion* because there may be other constructs that we did not measure that influence both the background and the pupils’ emotional state regarding school and language learning. Moreover, based on the literature on educational sociology, it is reasonable to assume that there is a cyclic relationship between background and school performance, that is, pupils’ skills and learning processes are not simply a consequence of individual differences, but parents’ and teachers’ expectations as well as other mechanisms of educational systems have a measurable impact on them (see Jussim & Harber 2005 on expectancy or Pygmalion effects).

Although they emerged as two independent factors in the analysis of Chapter 3, the construct of *academic emotion* is not completely orthogonal to the factor *cognition*. Thus, it is reasonable to allow for a correlation between these two latent constructs. Cognition, in turn, is again supposed to be correlated with the family background, for the same reasons that apply to the family background and *academic emotion*: As an example, we know from twin studies that achievement scores in foreign language, similarly to intelligence scores, can be explained in part by genetics (Rimfeld et al. 2015, Stromswold 2001). Thus, cognitive abilities and family background, while correlated, are not in a simple unidirectional causal relationship.

English proficiency at T1 may be influenced by the aforementioned constructs. The impact of the two factors *cognition* and *academic emotion* has been established in the analysis in Chapter 3. The factor *extrinsic motivation*, if *cognition* and *academic emotion* are controlled for, is not a meaningful predictor of English skills and is thus not included in the present analysis. Family background as well as the family’s linguistic repertoire (e.g., German as a family language) may both directly and indirectly (via the two aforementioned factors) explain variance in English at T1.

A first structural model is presented in Figure 3.⁴ In this model, I assume only one construct capturing the socioeconomic and educational background of the parents (*socioeco*). The variables capturing the first languages and the place where the child was born are added as manifest variables feeding into this construct. They arguably tap into relevant features of the family background: If the local language German is one of the family languages and if the child was born in Switzerland, then one can argue that the family is better lined up to conform with the educational system's expectations.

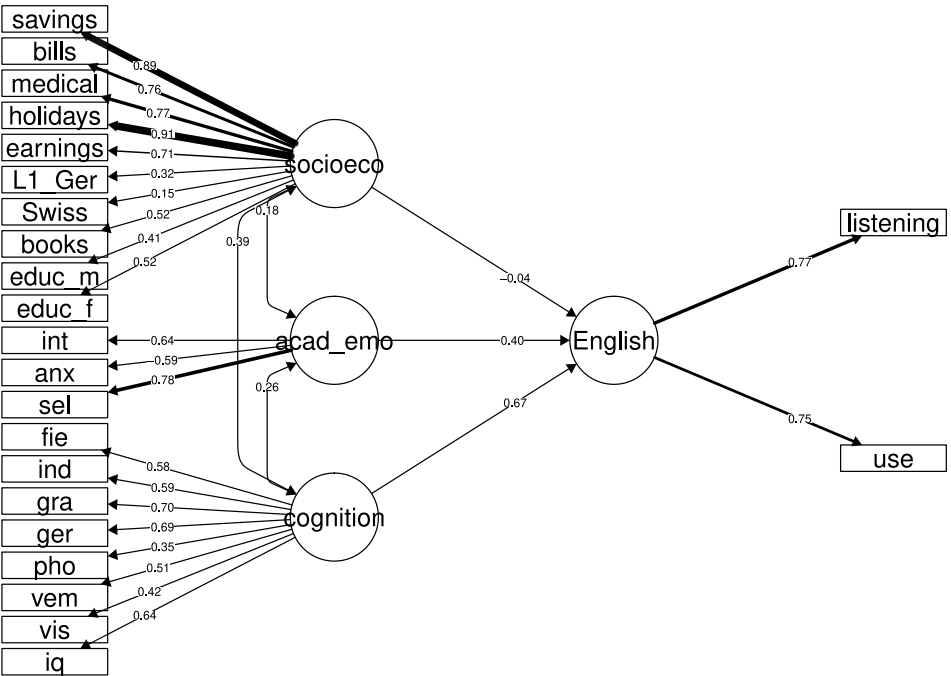


Figure 3: The constructs and their assumed associations; variant A with only one construct for family background (viz., “socioeco”).

Based on the literature it is reasonable to have a somewhat more nuanced view on family background. As argued in the seminal work by Bourdieu (1979), social inequalities are more than a simple matter of economic resources. As I have argued earlier in this chapter, sociologists of education in the wake of Bourdieu thus use at least a two-dimensional space to locate groups with respect to economic and cultural predispositions, and scoring high on one of the two does not

⁴More information on all models, including path plots with standardized estimates, can be found in the online supplementary materials.

imply a high score on the other. The next, more realistic model thus includes two correlated constructs: *cult_p* to capture cultural predispositions (high values expressing higher affinity and proximity to the established “high” culture) and *econ_p* to capture the economic resources. The first analysis in the next section will test whether assuming this more complicated model (Figure 4) is empirically justified.

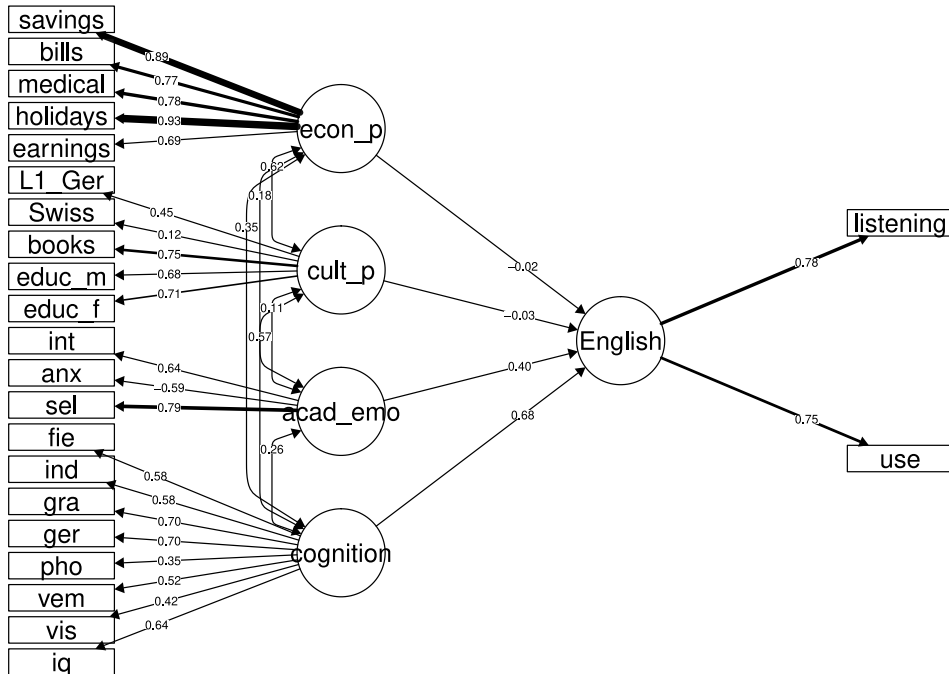


Figure 4: The constructs and their assumed associations; variant B with two constructs for family background (viz., *cult_p* and *econ_p*). The figure includes the standardized estimates (see next section).

In the structural equation models fitted in this chapter I do not include any correlations between the residuals of the manifest variables.

In the next section different variants of the model are fitted to the data, and the optimal solution is selected based on fit indices and model comparisons.

2.2 Socio-economic background: One or two dimensions

The first analysis investigates whether one or two constructs are needed to capture the family background and its impact on the other constructs. The two models specified in Figures 3 and 4 are fitted to the data using the `sem()` function using

the MLR estimator, which is suitable for clustered and incomplete data (Rosseel 2012). The full output of these models can be obtained from the supplementary materials. Table 2 compares these two models (and a few more models that will be discussed below) fitted in terms of three fit indices that are generally used to assess such models.

Table 2: Five different models fitted and three fit indices (root mean square error of approximation RMSEA, comparative fit index CFI, and standardized root mean residual SRMR).

Model variant	RMSEA	CFI	SRMR
A one construct (socioeco)	0.07	0.86	0.08
B two constructs (econ_p, cult_p)	0.05	0.93	0.06
C two constructs and regression with German as one of the native languages (L1_Ger)	0.05	0.93	0.06
D two constructs and regression with the status of being born in Switzerland or not (Swiss)	0.05	0.93	0.06
E two constructs, no Swiss and L1_Ger variables	0.06	0.93	0.06

In a first step, I compare a model with one background construction to the model with two constructs social and economic predispositions (models A and B in Table 2; see supplementary materials, Table S5.2 for full details). The comparison of the models is a way to quantify the trade-off between model simplicity and its goodness of fit: Adding more predictors will almost unavoidably increase the model fit, but it also increases the danger of overfitting that model to the specific data set collected by the researcher. Comparing two information criteria (Akaike information criterion, AIC and Bayesian information criterion, BIC) of the alternative models helps us decide which one is preferable. The model with two constructs (B) has more parameters to be estimated and fewer degrees of freedom than the smaller model (A). The smaller model is nested within the larger model. The model with two constructs has better values for the relative model quality (the AIC and BIC indices are both lower, by 222 and 206 respectively), and the χ^2 -test indicates that model with two latent constructs should be preferred (χ^2 difference 202, $p < 0.001$). We can therefore conclude that there is empirical support for the idea that different dimensions pertaining to the family background can be identified.

2.3 Family origin and home languages

After validating the two-dimensional nature of the family background, we address two other questions raised above: What is the influence of immigrant status and of other languages than German being spoken at home on the constructs under investigation? To assess these questions, the better model from the previous section is compared to two other models (see Figures 5 and 6).

The first comparison tests the prediction that not speaking German at home is an impediment to learning an additional language. The assumption is that pupils who need to develop their L2 German at the same time as the foreign language English are overwhelmed by the learning task. The new model (C in Table 2) in Figure 5 includes a regression arrow from the manifest variable L1_Ger to English proficiency. This allows us to assess the degree of linear association of the two variables if all other variables are controlled for.

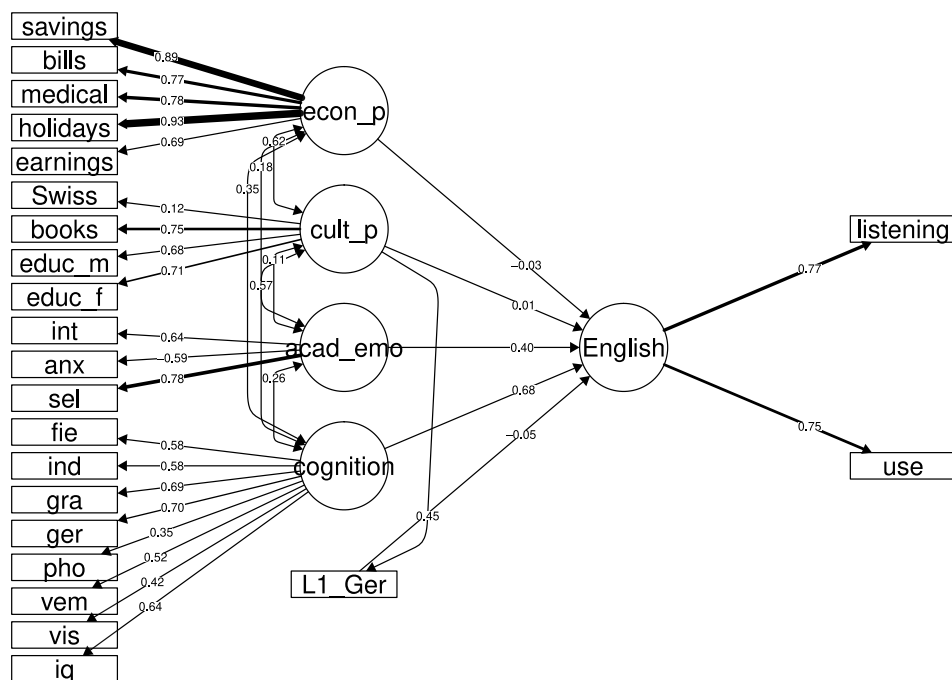


Figure 5: Model C fitted to test the impact of being an L1 German speaker or not. The Variable L1_Ger in this structural model affects English both directly and indirectly, first as a regressor and second via the latent construct of cultural predispositions.

The estimate for the regression from *L1_Ger* to English is very small (-0.052), and the model comparison shows that adding this parameter to the structural equation model is not justified: The additional parameter leads to roughly similar AIC (+1) and BIC (+5) values, and the χ^2 difference is negligible (0.89 , $p = 0.35$; see Table S5.3 in the supplementary materials for full details).

We conclude that, when controlling for all other variables in the model, speaking German in the family or not does not account for much variability in English as a foreign language skills.

The second analysis follows an analogous logic. Policymakers worry about the difficulty for foreign-born pupils to learn both the local language and foreign languages in primary school, in particular if their home languages are typologically distant from the (European) languages used and taught at school. It is therefore useful to add the *Swiss* variable (see Table 1) to test the impact of being born abroad or in Switzerland on English foreign language skills while controlling for all other variables.⁵ Figure 6 shows the structural model (model D in Table 2) that includes an additional parameter estimate for this background variable.

Again, the standardized estimate for this additional regression is minute (-0.002). Adding this parameter to the structural equation model is not justified (AIC for the model with the additional estimate decreases by 2, BIC by 6, χ^2 difference is 0, see Table S5.4 in the supplementary materials for full details).

The two variables *Swiss* and *L1_Ger* by themselves do not seem to be strongly associated with English skills. To assess how relevant they are to the overall model, I also fitted a simpler model similar to model A but that does not include these two variables. This last model (model E in Table 2) has only a slightly worse fit to the data.

2.4 Discussion

In the light of the comparisons above, we conclude that a model with two constructs for economic and cultural predispositions of the pupils' families and the two factors *cognition* and *academic emotion* fits the data well. Figure 5 shows its standardized estimates. These estimates do not change if one accounts for the clustered nature of the data using the *lavaan.survey* package as suggested by Oberski (2014). For more detail refer to Figure S5.8 in the supplementary material.

⁵Two other variants are possible and have also been tested: Both parents born abroad, one parent born abroad. All variants yield very similar results.

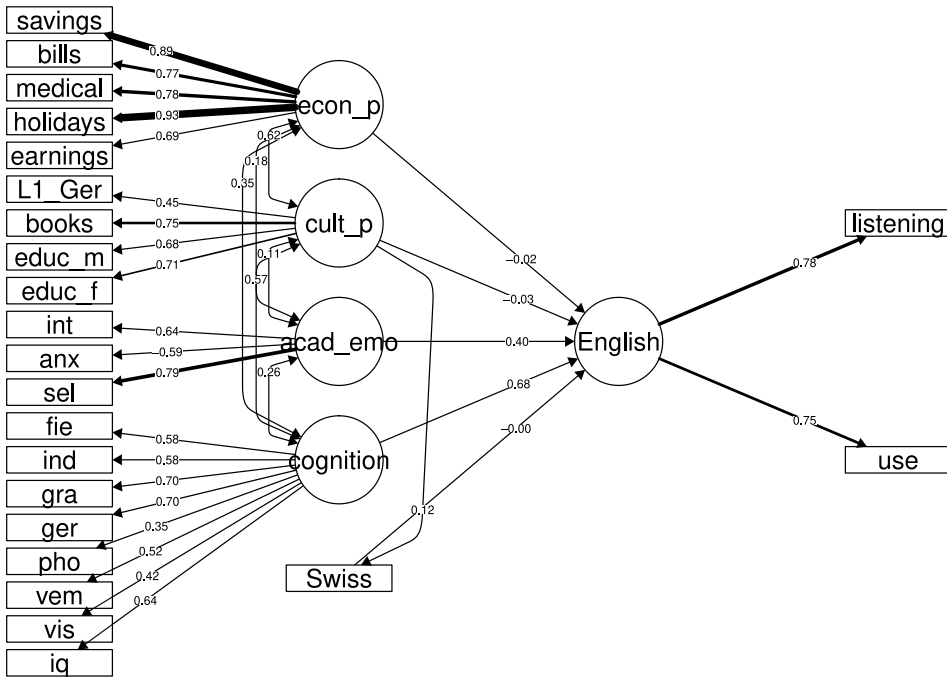


Figure 6: Adding the variable Swiss (pupil born in Switzerland or not) to the structural equation (Model D). The variable Swiss in this structural model affects English both directly and indirectly, first as a regressor and second via the latent construct of cultural predispositions.

The analysis shows, unsurprisingly in the light of the analyses in Chapter 3, that the two factors *cognition* and, to a lesser extent, *academic emotion*, are positively associated with English. Although the model allows the *economic* and *cultural* background constructs to directly influence English, the two corresponding estimates are very small (*cult_p*: -0.03, *econ_p*: -0.02). The impact of the two background constructs thus is predominantly indirect. English was not included in the *cult_p* construct since the overarching goal of our project is to investigate the favorable conditions and constraints on learning English as a foreign language. However, from a sociological point of view, language skills are undeniably part of the cultural resources of the families – which is the reason why German language background was included in the construct (Figure 5). The analyses in this chapter are in line with this construal of language being part of the family’s cultural resources: the associations from cultural background and cognition, the latter being in turn an important predictor of English skills, can also be analyzed in this light.

Table 3: Correlations of four constructs in model B.

factor	cult_p	econ_p
cognition	0.57	0.35
acad_emo	0.12	0.18

As shown in Figure 4 and Table 3, the strongest associations emerge between cognition and cultural predispositions. To a somewhat lesser extent, economic predispositions are also positively associated with cognition. The two constructs operationalizing two dimensions of the family background thus are first and foremost associated with the cognitive abilities of the pupils.

The literature on multilingual education often points out the incompatibility of the (often monolingual) *habitus* expected by Western educational systems with the attitudes of low-income and immigrant families. The term *habitus* is borrowed from the Bourdieu approach, and typically authors in the field of intercultural studies and multilingual education advocate a new, different, multilingual *habitus* that is hoped to make the educational systems a better place for minorities (Gogolin 1994). Based on such assumptions, one would expect the association between these background variables and the attitudinal and emotional construct (*acad_emo*) to be important: Low-income migrant families would arguably foster attitudes that clash with the educational system's expectations, and as a consequence children from such backgrounds should feel emotionally less at ease with respect to schooling and school subjects. In our data, it seems that this association is not particularly strong. This could either mean that expectations based on this *habitus* approach are not empirically borne out in the first place or that the schools under investigation have managed to overcome the differences implied by the *habitus*-based theory.

As a last sanity check preventing us from drawing hasty conclusions on the relative impact of being a migrant, we can fit the preferred model discussed here to different groups. As an example, the estimates can be compared across subgroups with one or two Swiss parents and the rest of the sample. Both variants yield very similar estimates across groups (see Figure S5.9. in the supplementary material). This supports the insight that once we account for the cognitive and affective individual differences as well as for the families' socioeconomic and educational background, having a migrant background or not is not distinctive when it comes to the learning of English as a foreign language at school.

3 Conclusion

In this chapter, an additional source of individual differences was added to the variables already discussed in the previous chapters. Based on information about the pupils' families' educational and economic characteristics and their home languages, different structural equation models were fitted to the data.

The comparison of the fitted models shows that there is indeed a sound empirical basis for the classic distinction between economic and cultural predispositions of the participants. The analysis shows furthermore that these two constructs, if all other factors are held constant, are not directly associated with the foreign language skills. They are, however, associated with the two factors discussed in Chapter 3 – most importantly with the construct that we termed *cognition*. This construct involves not only memory and intelligence, but also language-specific components such as grammatical sensitivity, inductive learning, etc. I deliberately modelled a two-way relationship between cognition and academic emotion on the one hand and cultural predispositions on the other, since we know that other constructs that were not measured have an impact on both. The path leading from cultural and economic backgrounds to English skills via cognition, however, raises a few questions. Not only, so it seems, are smart children⁶ obviously more likely to achieve better in the foreign language, but these smart children are more likely to come from families with high cultural interests (that are often also economically well-off). One of the widely shared expectations is that educational systems should somehow manage to even out these inequalities and create more “equality of chances”, which boils down to comparable or almost equal skill levels at crucial moments of institutional selection (see Heid 1988 for a critical discussion of this postulate). If the schools in which the pupils in our sample are educated produced more equality in this sense, the association between the background variables and constructs such as cognition

⁶I am aware that in the context of (language) pedagogy it is not fashionable to refer to smart and not-so-smart children. To some scholars in the field the mere idea of testing IQ, or sometimes testing any individual ability at all, seems problematic or scandalous (see Foucault 1975: 186 for an influential text in this respect). I certainly agree with the critique of abusive testing and the nefarious use of test scores in education (see Kuhn & Mai 2015 for an example on language test scores in primary school), research, and politics. At the same time, I do not think that such abuse necessarily means that testing abilities is intrinsically bad. Ignoring robust evidence on individual differences in cognitive abilities simply because they are not compatible with one's social and political ideologies is not only unscientific, but also problematic from the point of view of social policy. As Plomin (2019: chapter 9) argues, understanding individual differences, in particular if they are not caused by environmental biases, is not antithetical to the scholarly interest in equality and equal opportunity but on the contrary crucial for the assessment of such questions.

should be weak. But it is not. For T1, we diagnose an indirect association of family background on cognition and the emotional construct, which in turn has an impact on English skills. English skills at T1 furthermore are highly predictive of English skills at T3 (see Chapter 4). As far as our analyses allow any insight in the school's potential to even out socially caused inequalities, they seem to be neither reduced nor exacerbated by the system. The latter would be the case if, on top of their advantage already visible at T1, social variables explained additional variance beyond what is acquired and measured at T1 for the prediction of skills at T3. This, as Chapter 4 shows, is also not the case.

An impact of social background on school achievement would be completely unsurprising if the main object of investigation had been skills in mathematics or in the language of instruction, as was the case in Kronig (2003). However, according to the dominant view in multilingualism research (Cenoz 2003, Herdina & Jessner 2002, Montanari 2019), the subject we investigated is deemed to provide a headstart for a vulnerable subgroup of children: Immigrants who speak different languages at home should already have gathered useful experience in learning new languages and should benefit from a multilingual boost (see Berthele & Udry 2019 for references to studies with null or negative effects as well as an empirical investigation of this claim). It seems, at least in our data, that not even in a subject that should provide an advantage to pupils with an immigrant background such an effect can be found when controlling for the cultural and economic characteristics of the families.

Acknowledgments

I would like to thank Alexandre Duchêne, Jan Vanhove and three anonymous reviewers for their invaluable comments on earlier versions of this chapter.

References

- Avineri, Netta & Eric J. Johnson. 2015. Bridging the language gap. *Journal of Linguistic Anthropology* 25.
- Berthele, Raphael & Isabelle Udry. 2019. Multilingual boost vs. cognitive abilities: Testing two theories of multilingual language learning in a primary school context. *International Journal of Multilingualism*. 1–20. DOI: 10.1080/14790718.2019.1632315.
- Bourdieu, Pierre. 1979. *La distinction: Critique sociale du jugement*. Paris: Les Éditions de Minuit.

- Cenoz, Jasone. 2003. The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism* 7(1). 71–87. DOI: 10.1177/13670069030070010501.
- DESI-Konsortium. 2008. *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie*. Weinheim: Beltz.
- Entwisle, Doris R. & Karl L. Alexander. 1992. Summer setback: Race, poverty, school composition, and mathematics achievement in the first two years of school. *American Sociological Review* 57(1). 72–84. DOI: 10.2307/2096145.
- Farkas, George. 2018. Family, schooling, and cultural capital. In Barbara Schneider (ed.), *Handbook of the sociology of education in the 21st century*, 3–38. Cham: Springer. DOI: 10.1007/978-3-319-76694-2_1.
- Fernald, Anne, Virginia A. Marchman & Adriana Weisleder. 2013. SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science* 16(2). 234–248. DOI: 10.1111/desc.12019.
- Foucault, Michel. 1975. *Surveiller et punir: Naissance de la prison* (Bibliothèque des histoires). Paris: Gallimard.
- Gogolin, Ingrid. 1994. *Der monolinguale Habitus der multilingualen Schule*. Münster: Waxmann.
- Häberlin, Urs, Christian Imdorf & Winfried Kronig. 2005. Verzernte Chancen auf dem Lehrstellenmarkt: Untersuchungen zu Benachteiligungen von ausländischen und von weiblichen Jugendlichen bei der Suche nach beruflichen Ausbildungsplätzen in der Schweiz. *Zeitschrift für Pädagogik* 51.
- Heid, Helmut. 1988. Zur Paradoxie der bildungspolitischen Forderung nach Chancengleichheit. *Zeitschrift für Pädagogik* 34. 1–17.
- Herdina, Philip & Ulrike Jessner. 2002. *A dynamic model of multilingualism: Perspectives of change in psycholinguistics*. Clevedon: Multilingual Matters.
- Jussim, Lee & Kent D. Harber. 2005. Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review* 9(2). 131–155. DOI: 10.1207/s15327957pspr0902_3.
- Kronig, Winfried. 2003. Das Konstrukt des leistungsschwachen Migrantenkindes. *Zeitschrift für Erziehungswissenschaft* 6(1). 124–139.
- Kübler, Markus, Roland Kammer & Cordula Schneckenburger. 2014. *Fremdsprachenunterricht in der Volksschule: Ein Überblick über die Argumente und den Forschungsstand (Expertise)*. Schaffhausen: Lehrerinnen und Lehrer Schaffhausen.
- Kuhn, Melanie & Miriam Mai. 2015. Ethnisierende Differenzierungen im Elementarbereich: Sprachstandserhebungsverfahren zwischen landes- und kommunalpolitischen Direktiven und situiertem eigenlogischen Vollzug. In Thomas

- Geier & Karin Zabrowski (eds.), *Migration – Auflösungen und Grenzziehungen: Perspektiven einer erziehungswissenschaftlichen Migrationsforschung*, 115–134. Wiesbaden: VS-Verlag.
- Lucas, Samuel R. & Véronique Irwin. 2018. Race, class and theories of inequality in the sociology of education. In Barbara Schneider (ed.), *Handbook of the sociology of education in the 21st century* (Handbooks of Sociology and Social Research), 73–107. Springer. DOI: 10.1007/978-3-319-76694-2.
- Montanari, Simona. 2019. Facilitated language learning in multilinguals. In Simona Montanari & Suzanne Quay (eds.), *Multidisciplinary perspectives on multilingualism: The fundamentals*, 299–324. Berlin, Boston: De Gruyter Mouton. DOI: 10.1515/9781501507984.
- Oberski, Daniel. 2014. Lavaan.survey: An R package for complex survey analysis of structural equation models. *Journal of Statistical Software* 57(1). 1–27. DOI: 10.18637/jss.v057.i01.
- Plomin, Robert. 2019. *Blueprint, with a new afterword: How DNA makes us who we are*. London: Penguin Books.
- Riley, Dylan. 2017. Bourdieu's class theory. *The Catalyst* 1(2). 107–136.
- Rimfeld, Kaili, Philip S. Dale & Robert Plomin. 2015. How specific is second language-learning ability? A twin study exploring the contributions of first language achievement and intelligence to second language achievement. *Translational Psychiatry* 5(9). 638. DOI: 10.1177/0956797610368060.
- Rosseel, Yves. 2012. Lavaan: An R package for structural equation modeling. *Journal of Statistical Software* 48(2). 1–36.
- Slobin, Dan I. 1985. Crosslinguistic evidence for the language-making capacity. In Dan I. Slobin (ed.), *The crosslinguistic study of language acquisition*, 1999–1256. Hillsdale, N.J.: Lawrence Erlbaum.
- Stromswold, Karin. 2001. The heritability of language: A review and meta-analysis of twin, adoption, and linkage studies. *Language* 77(4). 647–723. DOI: 10.1353/lan.2001.0247.
- Vanhove, Jan. 2021. *Language aptitude in primary school*. Tech. rep. Fribourg, Switzerland. <https://osf.io/d9gnh/>.

Chapter 6

Creative thinking as an individual difference in task-based language teaching and learning

Isabelle Udry^{a,b}


^aUniversity of Fribourg, Institut de Plurilinguisme ^bZurich University of Teacher Education

Creative thinking is an individual difference variable worth to be investigated in the context of task-based language teaching (TBLT) which requires learners to accomplish communicative tasks by using their own ideas. It is hypothesized that creative students are at an advantage in the TBLT classroom because they produce and elaborate ideas more easily and have therefore more opportunity to engage with the target language. The present chapter seeks to clarify the role of creative thinking in instructed foreign language (FL) learning for primary school children who are taught in the TBLT paradigm. The research question is whether creative thinking has an effect on a) FL proficiency and b) FL motivation. To this aim, 87 learners of L2 French and L3 English (mean age 11.1 years) completed tests on non-verbal creative thinking, crystallized intelligence, C-Tests (L2 French), reading and listening comprehension (L3 English), as well as a questionnaire on FL motivation. Data was analyzed by means of structural equation modelling. The results show an association between creative thinking and FL proficiency. No effect was found for creative thinking on FL motivation.

1 Rationale for this chapter

This chapter deals with the cognitive mechanisms underlying creativity and their impact on foreign language (FL) proficiency and motivation. Task-based language teaching (TBLT) is the common method of instruction across Switzerland.



Isabelle Udry. 2021. Creative thinking as an individual difference in task-based language teaching and learning. In Raphael Berthele & Isabelle Udry (eds.), *Individual differences in early instructed language learning: The role of language aptitude, cognition, and motivation*, 125–141. Berlin: Language Science Press. DOI: 10.5281/zenodo.5464755 

TBLT conveys language through meaningful tasks that need to be completed using the target language (Willis 1996, Ellis 2017; see §2.3 for details). In the region of this study, TBLT often requires students to use language in connection with creative thinking to produce an outcome, such as a poem, a rap, or a role-play. It is reasonable to assume that creative individuals cope better with such tasks, because they are more skilled at generating and developing ideas. As a result, they may have more opportunity to engage with the target language (Ottó 1998, Albert 2006). From a language processing perspective, creative people are assumed to be better equipped to handle basic processes, such as filtering linguistic input, dealing with novelty, and tolerating ambiguity, thus potentially progressing faster at language learning (Kharkhurin 2012).

Furthermore, creative children may simply have more fun in the TBLT classroom because they can use their imagination and ideas in their production. This may foster their enjoyment of learning a foreign language and their intrinsic motivation to do so.

In terms of classroom teaching, some authors have speculated that if creativity could enhance language learning capacity, it might be a good idea to promote creative thinking with specific programs (Kharkhurin 2012). In the creative cognition rationale, such a training is deemed possible, even desirable in an educational context (Finke et al. 1992, Vogt 2010). At the same time, the needs of less creative learners would have to be considered carefully as they could be disadvantaged in a learning environment that strongly builds on creative thinking (Ottó 1998). Similarly, a reversed association can be assumed, i.e. that language learning has a favourable impact on creative thinking (Ghonsooly & Showqi 2012).

Dörnyei & Ryan (2015) suggest that creativity is a relevant ID for FL learning in current language teaching methodology. But rather than focusing on its relationship with other IDs, they argue for “paying greater attention to the interface between an individual’s inherent creativity as a predisposition and the external environment” (Dörnyei & Ryan 2015: 175). The present study adopts this perspective by focusing on the role of children’s creative thinking in the TBLT environment with a sample of 87 children (mean age 11.1) learning L2 French and L3 English from the project Language Aptitude at Primary School (LAPS) I.

2 Review of the literature

2.1 Creative thinking and foreign language learning

Creativity covers an array of abilities that are reflected just as much in outstanding artistic performances, scientific innovation, or everyday ingenuity. The construct subsumes cognitive, motivational, personality-linked, societal and proce-

dural aspects which have been incorporated into different theories of creativity (for an overview, see Lubart 1994). Cognitive mechanisms underlying creativity have been described as one of the construct's key components (Lubart 1994) and are thought to draw on the same cognitive structures as other intellectual abilities. Several authors have therefore integrated creativity into theories of intellect, namely Guilford's (1959) Structure of intellect, Sternberg's (1985) theory of successful intelligence, or Carroll's (1993) model of human cognitive abilities.

The relationship between creativity and intelligence has been discussed extensively (Sternberg & O'Hara 1999) and currently, they are considered separate, but overlapping constructs (Vogt 2010).

As early as 1959, Guilford described two distinct but complementary creative thought processes: Divergent thinking (ability to produce and elaborate many different and/or unusual ideas) and convergent thinking (ability to select good ideas and turn them into a product). They are still at the core of creative cognition studies today, with divergent thinking being more frequently investigated (Cropley 2006). This study therefore defines creative thinking as divergent thinking and the terms will be used synonymously in the remainder of the chapter.

Divergent thinking draws on memory functioning, executive control, and metacognition. More specifically, it requires an individual to cope with a new situation by retrieving existing knowledge, focusing on important information and transforming it in order to form novel associations and generate ideas to solve a problem or task. To do this successfully, individuals must separate relevant from irrelevant information and tolerate ambiguity when an answer is not immediately available (Finke et al. 1992, Cropley 2006).

Similar processes are hypothesised to be involved in learning a foreign language. For instance, Grigorenko et al. (2000) emphasize that successful language learners are able to deal well with novelty and tolerate ambiguity in the face of new and unfiltered linguistic material. Also, they can access existing knowledge easily and merge it with new information in order to fill linguistic gaps. People with these abilities are speculated to be at the same time creative and good language learners (Kharkhurin 2012).

2.2 Studies on creative thinking and L2 learning

An early study by Carroll (1964) mentioned in Ottó (1998) identified creativity as a poor predictor of L2 proficiency. In a review of personality correlates contributing to L2 achievement, Gardner (1990) even reports a negative correlation between creativity and French proficiency. The authors argue that these findings

may be linked to the audiolingual method prevalent at the time of the investigation which concentrated on teaching linguistic structures, leaving little scope for creativity (Ottó 1998, Gardner 1990). Therefore, Ottó (1998) readdressed the question in a more contemporary communicative teaching environment. He found significant positive correlations between creative thinking and L2 English grades in 36 Hungarian high school students. These results could not be replicated by Albert (2006) who investigated the relationship between L2-proficiency, creative thinking and language aptitude in 41 advanced L2 English learners in their first year at university. The author even found negative correlations between some aspects of creative thinking (fluency and flexibility) and phonetic coding ability which is a component of language aptitude. Albert & Kormos (2011) revisited the issue by focusing on narration and creativity in 76 Hungarian English learners at secondary school (15–16 years). Again, they observed no significant correlation between creativity and narration complexity, accuracy and lexical diversity.

The influence of language tuition on creative thinking was examined by (Ghonsooly & Showqi 2012). They compared creative thinking in advanced English learners ($n = 60$, 15–16 years) to a control group ($n = 60$) of the same age that had never attended foreign language classes. Intelligence was controlled for with a non-verbal intelligence test. English learners scored significantly higher on creative thinking than the control group. In a follow up study (Showqi & Ghonsooly 2015) compared language awareness and creative thinking in beginning and advanced English learners. In both groups, significant positive correlations between the two constructs were identified. Based on their results, the authors argue that language tuition is beneficial for creative thinking.

Fleith et al. (2002) examined the effects of a creativity training intervention on divergent thinking in Brazilian children who attended either a monolingual Portuguese school or a bilingual Portuguese-English program. Their aim was twofold: to assess training effects from the intervention and investigate group differences between immersion and conventional students. Treatment slightly improved divergent thinking abilities. However, divergent thinking was not related to attendance of a mono- or bilingual school program.

Studies reported here have produced mixed findings. This may be due in part to methodological choices, i.e the fact that different age groups and aspects of language were investigated, making it difficult to compare results.

2.3 Creative thinking and TBLT

Creative thinking plays a role in educational achievement, especially when knowledge is conveyed through open-ended, game like elements that incite ac-

tive involvement (Runco 2004). As far as foreign language learning is concerned, TBLT is a method that can meet these criteria. Tasks can be defined as goal-oriented, meaning-focused activities which require learners to use language in order to achieve a real outcome (Willis 1996: 2). In addition, tasks are meaningful, i.e. they relate to real-life, and contain some kind of a gap that needs to be filled with information or an opinion (Ellis 2017). Tasks can be focused (with a communicative aim using a particular language structure), or unfocused (with a general communicative aim) (Ellis 2009).

TBLT has developed from communicative language teaching (CLT, Krashen 1981) and, over time, a weak and strong form have emerged (Ellis 2017: 109). In its strong form, TBLT aims at learning *through* communication independent of curricular prescriptions (Ellis 2017). Tasks guide the learning process as they provide the medium by which structural issues are raised, for instance when a specific communicative need or problem arises. The correct use of the target language is secondary.

The weak form of TBLT is found in task-supported language teaching (TSLT) which, like its *strong* counterpart, primarily fosters communicative skills in the target language. Contrary to the strong form, TSLT draws on a prescribed syllabus and language is conveyed in a way that resembles the presentation-practice-production (PPP) model. In TSLT, tasks are considered an opportunity for learners to use the language they have learnt in a communicative context, usually in the production phase of the teaching cycle. In terms of vocabulary and grammar teaching, TSLT directs learner's attention to specific linguistic forms, often explicitly in a focus-on-form element of a lesson. While communication is clearly the main objective of TSLT, the correct use of the target language is also emphasized to some extent.

TBLT is implemented differently by teachers (Ellis 2009) and there is a variety of task types that draw on different learner skills and abilities (Skehan 1996). For a discussion of key concepts and common misunderstandings relating to TBLT, the reader is referred to Ellis (2009).

2.4 TBLT in the Swiss curriculum

As detailed in Chapter 2 this volume, the Swiss curriculum emphasizes communication in the target language and places less importance on the correct use of form. Teaching methods in the region of this study are oriented towards CLT and more specifically a weak form of TBLT. The teaching manuals, which are pre-

scribed by the local boards of education,¹ are structured around teaching units on specific topics (e.g., sports, art, or storytelling) which are introduced to learners via authentic written or auditory input. The authentic input is consolidated with a series of meaning-based activities that prepare pupils for the communicative task at the end of each unit. In the process, pupils are made familiar with some predetermined aspects of grammar and vocabulary.

The teaching manuals highlight the use of creative thinking in the communicative task where learners need to draw on the target language to plan and elaborate a creative product. For instance, children design a poster about a particular sport, create a piece of artwork and describe it to their peers at a vernissage, or turn a short story into a play that is presented to the class (Arnet-Clark et al. 2013). Group presentations are followed up with a personal reflection on the task process. “Project tasks”, as they are called in the L3 English manual (Arnet-Clark et al. 2013), are carried out individually or in groups. The hypothesized link with creative thinking outlined in §2.1 is particularly evident in this type of TBLT scenario. Individual differences in creative thinking are therefore expected to affect task performance and consequently language proficiency.

2.5 Creative thinking and FL motivation

FL motivation and related constructs are discussed in Chapter 1, §4. The link between creative thinking and FL motivation in the TBLT classroom is underpinned by Deci & Ryan’s (1985, 2002) self-determination theory (SDT). A key concept of SDT is intrinsic motivation, i.e. an inherent drive to engage with the environment that is largely independent of external sources. Intrinsic motivation resides on three basic individual needs: self-determination, competence, and interpersonal relatedness. I argue that TBLT with a focus on creative thinking can meet these needs in several ways: giving children the opportunity to elaborate on their own ideas contributes to their feeling autonomous and competent. Working with others and getting feedback on their outcomes strengthens a sense of competence and connectedness to the group.

The opportunity to be creative while studying a language is expected to affect enjoyment of learning and therefore intrinsic FL motivation. Forms of extrinsic motivation, such as wanting to earn good grades or keep up with other students, are deemed less relevant for this context and are not considered in the analysis. In what follows, *French motivation* and *English motivation* therefore refers to intrinsic motivation only.

¹In LAPS I schools: *Mille feuilles* (Bertschy et al. 2012) for L2 French; *New World* (Arnet-Clark et al. 2013) for L3 English

In terms of the structure of FL motivation as assessed in our student questionnaire, the factor analysis reported in Chapter 3 points to two different factors: Intrinsic motivation, L2 self-concept, and L2 anxiety form a common factor we call *L2 Academic Emotion*. Extrinsic items, dedication, and teacher/parental support load onto a separate factor we refer to as *Extrinsic Factor*. Regression analysis shows a substantial contribution of L2 Academic Emotion to L2 proficiency, while the Extrinsic Factor is not associated positively with L2 proficiency.

3 Method

3.1 Research questions

Two research questions are addressed in this chapter:

1. What is the association between creative (divergent) thinking and FL proficiency?

Based on current research findings and theoretical arguments on creative cognition and TBLT, it is hypothesized that creative thinking (assessed in a non-verbal test for divergent thinking) is positively related to children's FL proficiency. Learners with well-developed creative thinking skills are expected to be at an advantage in TBLT due to an enhanced ability to generate and elaborate ideas. It is assumed that they have more capacity to invest in actual language work, thus progressing faster at language learning.

2. Does creative thinking relate to students' intrinsic FL motivation in the TBLT classroom?

To the author's knowledge, this question has not been studied explicitly yet. Considering aspects of TBLT outlined in §2.3 and theories of FL motivation discussed in Chapter 1, a positive association between creative thinking and intrinsic motivation to learn in the TBLT classroom is hypothesized. TBLT is thought to enhance creative children's enjoyment of FL learning because particular attention is given to their ideas.

3.2 Participants and procedure

The research design of the LAPS project is described in Chapter 2, and only aspects relevant to this study are summarized here.

The sample has been derived from LAPS I and included a total of 117 5th graders (mean age 11.1 years in spring 2017) learning L2 French and L3 English. Children with L1 English or L1 French, as well as incomplete observations were excluded from the sample, leaving 87 complete observations for data analysis.

At T1 (spring 2017), creative thinking, intelligence, and L2 French were tested and children filled in a L2 French motivation questionnaire. A L2/L3 motivation questionnaire and L3 English test were administered a year later at T2. At the time of testing, participants had received 248 French lessons over 2.5 academic years (T1) and 152 English lessons over 1.5 academic years (T2).

Creativity was operationalized as divergent creative thinking and assessed with the Test of Creative Thinking (Divergent Production) (TCT-DP) by Urban & Jellen (1995). Participants need to complete an unfinished picture with 6 figural fragments presented to them on paper. There is no time constraint and participants are told that they are free to draw whatever they like. The test is scored according to 14 criteria (see Chapter 2, §3). This non-verbal test was chosen in order to rule out language as a confounding variable (Simonton 2008).

Intelligence was measured with the number sequencing subtest of the CFT 20-R (Weiß 2006) for crystallized intelligence.

L2 French proficiency was assessed with a C-Test composed of four short and independent paragraphs (Grotjahn 2002). L3 English proficiency was measured with the Oxford young learners' placement test OYLPT (Oxford English Testing 2013). It covers reading and listening comprehension. The test assesses language use, vocabulary, and grammar embedded in everyday situations. Due to the overall aims of the project, different skills were measured at different times (see Chapter 2 for details).

The motivational questionnaire is also detailed in Chapter 2. It covers various aspects, including intrinsic and extrinsic motivation, L2/L3 self-concept, L2/L3 anxiety, dedication, and perceived encouragement by teachers and parents. For reasons discussed in §2.5, *French motivation* and *English motivation* exclusively refers to items assessing intrinsic motivation.

3.3 Analysis

This chapter explores creative thinking in two ways: In relation to FL proficiency and intrinsic FL motivation. As outlined in §2.1, the literature postulates an overlap between intelligence and creative thinking. The two constructs may share some of the explained variance in the statistical model, leading to biased parameter estimates (overestimation of creative thinking effects). This can be accounted for by treating intelligence as a confounder and controlling for it in the analysis

(Rohrer 2018). As one of the reviewers pointed out, the assumed overlap may be more related to fluid, rather than crystallized intelligence and it may have been more revealing to administer a test of fluid intelligence. Due to the overall design of the project, however, this was not possible in LAPS I (see Chapter 2 for details).

Like all psychometric constructs, intelligence (and creativity for that matter) cannot be observed directly but only measured approximately by a psychometric test. When controlling for intelligence, we actually control for the observation of the construct expressed in the individual test score. The more precise this measure, the more accurate the estimates from the statistical model will be (Vanhove 2015). However, test scores are always obtained with some measurement error owed to test conditions or to the test instrument. Accounting for this measurement error in the statistical model will increase its accuracy (Brunner & Austin 2009). As suggested by Westfall & Yarkoni (2016), measurement error can be accounted for by calculating the error variance of a given variable based on the reliability coefficient of its test and including the error variance in the statistical model. When addressing construct-related issues, such as the ones in this chapter, the authors recommend using structural equation modelling (SEM) or other latent-variable approaches which allow for introducing measurement error directly to the statistical model. SEM refers to a group of statistical techniques that are commonly used in social sciences to investigate unobservable latent constructs, such as creativity or intelligence. In the structural equation model, these latent constructs are represented by one or more observed variables, i.e. the instruments that assess the unobservable construct. The relationship between the latent constructs can then be estimated with independent regression equations.

4 Results

4.1 Creative thinking and FL proficiency

Based on the arguments outlined in §3.3, SEM was used with the lavaan package for R (Rosseel 2012) to allow for a) investigation of latent constructs and b) consideration of measurement error to increase model accuracy.

For the first research question, FL proficiency (L2 French and L3 English) was introduced to the model as an endogenous (or outcome) variable. Based on the assumption of an underlying aptitude for foreign language learning, the two languages were treated as one latent variable. The choice is further based on the multicompetence framework which hypothesizes the integration of all languages known to the learner in a common linguistic repertoire (for a recent discussion

see Cook 2012). Creative thinking (TCT-DP) and intelligence (number sequencing CFT-20 R) were introduced to the model as exogenous (or predictor) variables.

In Figure 1, circles show latent variables (or constructs) and squares show measured variables. One headed arrows indicate the regression coefficients or parameter estimates, i.e. the strength of association between the variables. Double headed arrows stand for variances.

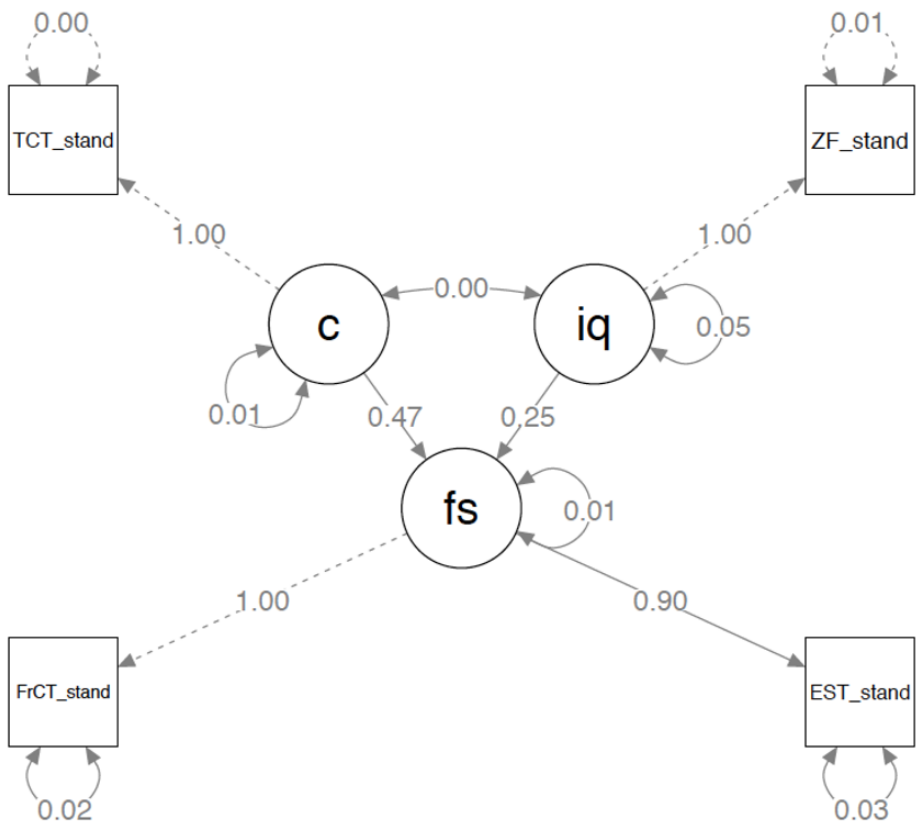


Figure 1: Path diagram for Model 3 (reliability 0.78). Nodes (latent constructs): fs = foreign language proficiency, c = creative thinking, iq = intelligence; TCT_stand = test on divergent thinking, ZF_stand = number sequencing test, FrCT_stand = French C-Test, EST_stand = English OYLPT score.

In order to determine measurement error, the reliability coefficients for the exogenous (predictor) variables (creative thinking, intelligence) were taken from

the test manuals to calculate error variances (Table 1). The obtained error variances were added to the different statistical Models 1–3.

For the number sequencing test, a reliability coefficient of Cronbach’s α 0.91 is reported (Weiß 2006). For the TCT-DP, different values are given, ranging from 0.38 to 0.78, depending on the validation study (Urban & Jellen 1995). Based on recommendations by Westfall & Yarkoni (2016), I accounted for the TCT-DP variation by fitting three models assuming reliability coefficients of 0.38 (Model 1), 0.58 (Model 2) and 0.78 (Model 3). For ease of reading, only Model 3 will be reported in the following. Further information on all models can be found at <https://osf.io/vxr9m/>.

Table 1: Exogenous (predictor) variables and their sample variance (Var), standard deviation (SD), reported reliability coefficients (RC) and calculated measurement error/error variance (ME) for this sample.

Model 3	Var	SD	RC	ME
Creative thinking (divergent thinking)	0.01	0.12	0.78	0.003
Intelligence (number sequencing)	0.06	0.25	0.91	0.006

Model 3 produced acceptable fit indices² (CFI = 1, RMSEA = 0, SRMR = 0.01). As shown in Table 2 and Figure 1, Model 3 supports an effect for creative thinking on language proficiency when intelligence is controlled for in the range of statistical significance. This result does not necessarily point to a direct causal influence from creative thinking to FL proficiency. The design used in this study does not allow for making claims on causality.

Table 2: Estimates for Model 3 (0.78) on the association between creative thinking and intelligence and FL proficiency.

L2/L3 proficiency ~	Est.	SE	<i>z</i>	<i>p</i>	ci lower	ci upper
Creative thinking	0.47	0.19	2.47	0.013	0.1	0.84
Intelligence	0.25	0.08	3.1	0.002	0.09	0.41

4.2 Creative thinking and FL motivation

The opportunity to be creative in the classroom is assumed to impact particularly on intrinsic FL motivation, rather than extrinsic motivation. The SEM for

²Cut off for good fit: CFI > 0.9; RMSEA < 0.08, SRMR < 0.08 (Kline 2011).

the second research question therefore includes intrinsic motivation for English and French as the endogenous (outcome) variable and creative thinking as the exogenous (predictor) variable. Different models with combinations of motivation items were fitted. A total of four items were retained in the final Model 4 represented in Figure 2 (the same symbols apply as in Figure 1). Combinations of other FL motivation items fitted the data less well and were therefore not pursued further.³

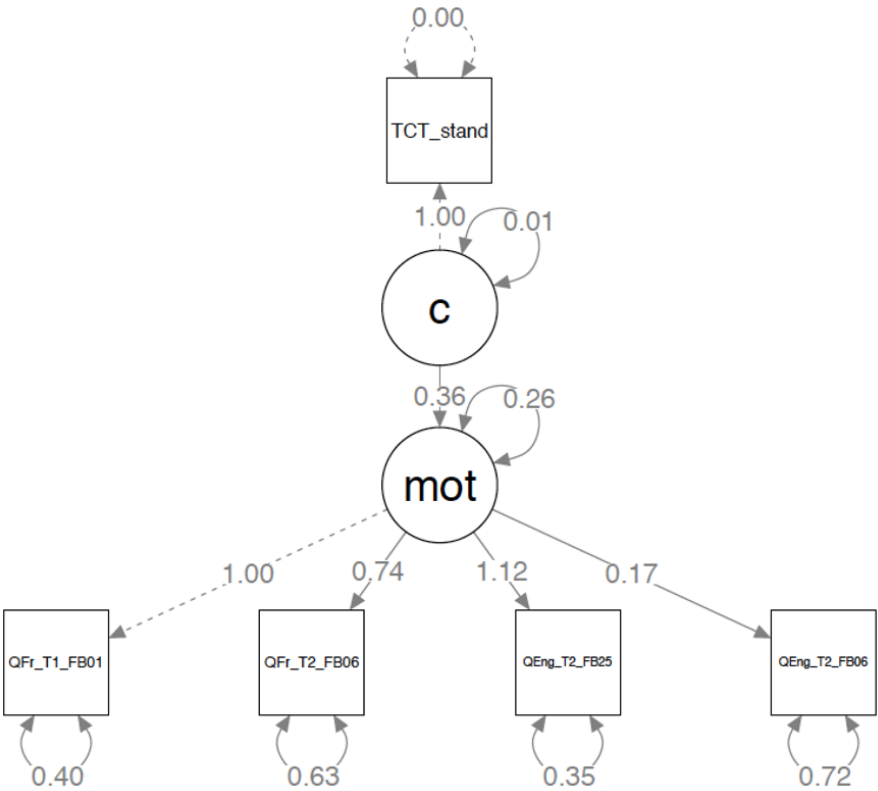


Figure 2: Path diagram for Model 4 (reliability 0.78). Nodes: c = creative thinking, mot = intrinsic motivation (latent constructs); TCT1 = test on divergent thinking, intrinsic motivation French: QFr_T1_FB01, QFr_T2_FB06, intrinsic motivation English: QEng_T2_FB25_QEng_T2_FB06.

³Kolenikov & Bollen (2012) describe several possible causes for unusual model indications, referred to as Heywood cases: Outliers, empirical underidentification, structural misspecification, missing data or sampling fluctuations. The authors also give an overview of how to address these issues.

Again, different measurement errors were considered for the creative thinking test in Model 4, all options yielded the same acceptable fit to the data (CFI = 1, SRMR = 0.04, RMSR = 0). The association between creative thinking and intrinsic motivation turns out to be negligible and non-significant, as indicated by the low z -value and high p -value reported in Table 3.

Table 3: Estimates from Model 4 (0.78) on the association between motivation to learn a foreign language and creative thinking.

Intrinsic Motivation ~	Est.	SE	z	p	ci lower	ci upper
Creative thinking	0.36	0.74	0.49	0.63	-1.1	1.81

5 Discussion

Research question 1 addressed the association between creative (divergent) thinking and FL proficiency. A statistically significant effect emerges from the data, indicating that creative thinking plays a role in children's developing FL proficiency when they are taught in the TBLT paradigm. The present study thus mirrors findings from Ottó (1998) with high school students and contradicts Albert (2006) and Albert & Kormos (2011) who could not replicate results from the Ottó study. Research question 2 explored the possibility that creative children are more motivated to learn foreign languages with TBLT than their peers, i.e. that high scores in the creative thinking test are associated with high values on the motivation-questionnaire items. This hypothesis could not be substantiated: the association between creative thinking and FL motivation in this sample is negligible and non-significant.

It is worth pointing out that the results reported here do not allow for stipulating any causal links between the investigated constructs. While an effect of creative thinking on FL proficiency has been found in the data, the direction of causality remains unclear. It may well be, as some scholars suggest, that language learning contributes to creative thinking. Such claims have been made mainly with reference to simultaneous bilingualism (for an overview, see Ricciardelli 1992), rather than exposure to instructed language learning. To address causality, Simonton (2008) suggests a longitudinal design with multiple assessment of the variables where time would allow for comparison within and between subjects. If multilingualism resulting from instructed language learning were a predictor for creative thinking, this would be detected in the data if an individual's

FL proficiency at T1 and creative thinking at T2 were more strongly correlated than creative thinking at T1 and individual FL proficiency at T2 (Simonton 2008: 154). However, this kind of research is rare and obviously the research design presented in this chapter does not allow for such inferences.

Some changes to the design could have improved the robustness of the findings. For instance, children's motivation to learn foreign languages did not include questions on how they liked the teaching methods or textbooks. Also, assessing creative thinking more comprehensively, including a range of tests and information on creative hobbies might have provided a more detailed view of the creative student than a mere non-verbal test. These aspects may be considered in future studies to provide further insights into the role of creative thinking in instructed language learning.

References

- Albert, Ágnes. 2006. Learner creativity as a potentially important individual variable: Examining the relationships between learner creativity, language aptitude and level of proficiency. *University of Pécs Roundtable 2006: Empirical Studies in English Applied Linguistics*. 77. <http://mek.niif.hu/04900/04902/04902.pdf#page=72> (15 April, 2017).
- Albert, Ágnes & Judit Kormos. 2011. Creativity and narrative task performance: An exploratory study. *Language Learning* 61(s1). 73–99.
- Arnet-Clark, Illya, Silvia Frank Schmid, Linda Grimes, Guido Ritter, Jean Rüdiger-Harper, Sandra Thum-Widmer & Chantal Villiger. 2013. *New World: English as a second foreign language*. Baar: Klett und Balmer.
- Bertschy, Ida, Barbara Grossenbacher, Esther Sauer & Andi Thommen. 2012. *Mille feuilles*. Bern: Schulverlag plus AG.
- Brunner, Jerry & Peter C. Austin. 2009. Inflation of Type I error rate in multiple regression when independent variables are measured with error. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique* 37(1). 33–46. <https://www.jstor.org/stable/25653459> (18 April, 2019).
- Carroll, John B. 1964. The prediction of success in intensive foreign language training. In *Training research and education*, 87–136. Pittsburgh: University of Pittsburgh Press.
- Carroll, John B. 1993. Abilities in the domain of idea production. In *Human cognitive abilities: A survey of factor analytic studies*, 394–439. Cambridge: Cambridge University Press.

- Cook, Vivian J. 2012. Multicompetence. In Carol A. Chapelle (ed.), *The encyclopedia of applied linguistics*, 1–6. Hoboken, NJ, USA: Blackwell Publishing. DOI: 10.1002/9781405198431.wbeal0778.
- Cropley, Arthur. 2006. In praise of convergent thinking. *Creativity Research Journal* 18(3). 391–404.
- Deci, Edward L. & Richard M. Ryan. 1985. *Intrinsic motivation and self-determination in human behaviour*. New York: Plenum.
- Deci, Edward L. & Richard M. Ryan (eds.). 2002. *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.
- Dörnyei, Zoltán & Stephen Ryan. 2015. *The psychology of the language learner revisited*. New York: Routledge.
- Ellis, Rod. 2009. Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics* 19(3). 221–246. DOI: 10.1111/j.1473-4192.2009.00231.x.
- Ellis, Rod. 2017. Task-based language teaching. In Shawn Loewen & Masatoshi Sato (eds.), *The Routledge handbook of instructed second language acquisition*, 108–125. New York: Routledge.
- Finke, Ronald A., Thomas B. Ward & Steven M. Smith. 1992. *Creative cognition*. Massachusetts: MIT Press.
- Fleith, Denise de Souza, Joseph S. Renzulli & Karen L. Westberg. 2002. Effects of a creativity training program on divergent thinking abilities and self-concept in monolingual and bilingual classrooms. *Creativity Research Journal* 14(3-4). 373–386.
- Gardner, Robert C. 1990. Attitudes, motivation, and personality as predictors of success in foreign language learning. In Thomas S. Parry & Charles W. Stansfield (eds.), *Language aptitude reconsidered*. Washington, D.C.: ERIC Clearinghouse.
- Ghonsooly, Behzad & Sara Showqi. 2012. The effects of foreign language learning on creativity. *English Language Teaching* 5(4). 161. DOI: 10.5539/elt.v5n4p161.
- Grigorenko, Elena L., Robert J. Sternberg & Madeline E. Ehrman. 2000. A theory-based approach to the measurement of foreign language learning ability: The Canal-F theory and test. *The Modern Language Journal* 84(3). 390–405.
- Grotjahn, Rüdiger. 2002. Konstruktion und Einsatz von C-Tests: Ein Leitfaden für die Praxis. In Rüdiger Grotjahn (ed.), *Der C-Test. Theoretische Grundlagen und praktische Anwendung*, vol. 4, 211–255. Bochum: AKS-Verlag.
- Guilford, Joy Paul. 1959. Three faces of intellect. *American Psychologist* 14. 469–479.
- Kharkhurin, Anatoliy V. 2012. *Multilingualism and creativity*. Bristol: Multilingual Matters.

- Kline, Rex B. 2011. *Principles and practice of structural equation modeling*. 3rd edn. New York: The Guilford Press.
- Kolenikov, Stanislav & Kenneth A. Bollen. 2012. Testing negative error variances: Is a Heywood case a symptom of misspecification? *Sociological Methods & Research* 41(1). 124–167.
- Krashen, Stephen. 1981. *Second language acquisition and second language learning*. Oxford: Pergamon Press. http://sdkrashen.com/content/books/sl_acquisition_and_learning.pdf (6 July, 2018).
- Lubart, Todd I. 1994. Creativity. In Robert Sternberg (ed.), *Handbook of perception and cognition. Thinking and problem solving*, 290–323. Cambridge, Massachusetts: Academic Press.
- Ottó, István. 1998. The relationship between individual differences in learner creativity and language learning success. *TESOL quarterly* 32(4). 763–773. <http://www.jstor.org/stable/3588011> (12 May, 2017).
- Oxford English Testing. 2013. *Oxford young learners placement test*. <https://www.oxfordenglishtesting.com/> (5 May, 2017).
- Ricciardelli, Lina A. 1992. Creativity and bilingualism. *The Journal of Creative Behavior* 26(4). 242–254. DOI: 10.1002/j.2162-6057.1992.tb01183.x.
- Rohrer, Julia M. 2018. Thinking clearly about correlations and causation: Graphical causal models for observational data. *Association for Psychological Science* 1(1). 27–42. DOI: 10.1177/251524591774562.
- Rosseel, Yves. 2012. Lavaan: An R package for structural equation modeling. *Journal of Statistical Software* 48(2). 1–36.
- Runco, Mark A. 2004. Creativity. *Annual Review of Psychology* 55. 657–687. <http://www.annualreviews.org/toc/psych/55/1> (12 May, 2017).
- Showqi, Sara & Behzad Ghonsooly. 2015. The Nature of Metacognitive Awareness in Foreign Language Learners and Its Interaction with Creativity. *Journal of Language Teaching and Research* 6.
- Simonton, Dean K. 2008. Bilingualism and creativity. In *An introduction to bilingualism principles and processes*, 147–166. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Skehan, Peter. 1996. A framework for the implementation of task-based instruction. *Applied Linguistics* 17(1). 38–62. DOI: 10.1093/applin/17.1.38.
- Sternberg, Robert J. 1985. *Beyond IQ: A triarchic theory of human intelligence*. Cambridge: CUP Archive.
- Sternberg, Robert J. & Linda A. O'Hara. 1999. Creativity and intelligence. In Robert J. Sternberg (ed.), *Handbook of Creativity*, 251–272. Cambridge: Cambridge University Press.

- Urban, Klaus K. & Hans G. Jellen. 1995. *Test zum Schöpferischen Denken – Zeichnerisch*. Frankfurt am Main: Pearson PsychCorp.
- Vanhove, Jan. 2015. *Controlling for confounding variables in correlational research: Four caveats*. <https://janhove.github.io/design/2015/08/24/caveats-confounds-correlational-designs> (28 April, 2020).
- Vogt, Thomas. 2010. Kreativität und Kognition. In *Kalkulierte Kreativität*, 189–231. VS Verlag für Sozialwissenschaften. DOI: 10.1007/978-3-531-92340-6_9.
- Weiß, Rudolf H. 2006. *Manual Grundintelligenztest Skala 2 – Revision (CFT 20-R)*. Göttingen: Hogrefe.
- Westfall, Jacob & Tal Yarkoni. 2016. Statistically controlling for confounding constructs is harder than you think. *PLOS One* 11(3). DOI: 10.1371/journal.pone.0152719.
- Willis, Jane. 1996. A flexible framework for task-based learning. In David Willis & Jane Willis (eds.), *Challenge and change in language teaching*, 52–62. New York: Macmillan Education.

Chapter 7

The closer the better? Investigating L2 motivation of young learners in different contexts

Carina Steiner


University of Bern, Center for the Study of Language and Society

One of the most prominent ID variables researched along with foreign language aptitude is L2 motivation. This chapter aims to provide insights in what constitutes L2 motivation of young learners who learn French and English as part of their mandatory curriculum. It includes two groups of primary school students from different regions in Switzerland, one living close to the French-German language border (LAPS I) and one far away (LAPS II). The effect of proximity to the French language border is of particular interest, since living close to the L2 speech community has been discussed as one of the most crucial influences on L2 motivation. Our results suggest that, while children are generally motivated to learn foreign languages at school, English is clearly favoured by both learner groups. Contrary to previous findings, our data does not corroborate a positive effect of proximity to the French language border on motivation to learn French. In turn, perceived encouragement and support given by the teacher seem to play a vital role in strengthening the learners' motivation, while effects of parental encouragement, gender and multilingual background are comparatively low.

1 Introduction

The LAPS project aims at understanding what shapes young learners' foreign language learning abilities, including a broad range of factors in addition to Carroll's (1958, 1964) traditional aptitude components (cf. Chapter 1, this volume). One of the most prominent factors in this context is L2 motivation. Whereas its interplay with other ID variables and its effect on language learning success was



Carina Steiner. 2021. The closer the better? Investigating L2 motivation of young learners in different contexts. In Raphael Berthele & Isabelle Udry (eds.), *Individual differences in early instructed language learning: The role of language aptitude, cognition, and motivation*, 143–161. Berlin: Language Science Press. DOI: 10.5281/zenodo.5464757 

subject of Chapter 3, this Chapter aims at a more detailed understanding of what constitutes L2 motivation itself.

In L2 research, motivation has been investigated from the late 50s onwards and is considered one of the most important factors explaining individual differences in language learning. Today, the vast number of publications in this field can hardly be captured, and yet, empirical studies focusing on young learners are comparatively rare. In addition, most of the existing analyses stay on a descriptive or correlational level and, to the author's knowledge, no direct comparison of foreign language learners in different sociocultural contexts has been made. The current study seeks to bridge this gap by investigating L2 motivation of two different groups of young learners in Switzerland, one living close to the French-German language border and one far away.

2 L2 motivation

Research and theory development in L2 motivation began with the pioneering work by Robert Gardner and Wallace Lambert in the late 1950s (Gardner & Lambert 1959). Based on extensive empirical research, they conceptualised L2 motivation under three essential aspects: interest in learning a language, attitudes towards the specific L2 and its speakers, and dedication towards learning the language (Gardner & Lambert 1959; see also Gardner 1985). The original model was later extended and reconceptualised in various ways, ranging from embedding general psychological theories (e.g. Noels 2001) to entirely new framings of theoretical core assumptions (e.g. Dörnyei 2005). Of particular importance are Deci & Ryan's (1985; cf. also Ryan & Deci 2002) Self-Determination Theory (SDT) with their central construct of intrinsic motivation and Dörnyei's (2005) L2 Motivational Self System (L2MSS) based on the idea that future projections of oneself affect motivation and, consequently, the L2 learning process. For a more detailed discussion on L2 motivation theories and research, the reader is referred to Chapter 1, §4, this volume.

In the current chapter, proximity to the L2 community is of particular interest. The basic goal of foreign language learning is to gain linguistic abilities that allow for active interaction with other people. Contact with L2 speakers permits learners to use their communicative language skills and reflect on them. Thus, emphasising the relevance of the L2 for real communication can become a driving force in the learning process.

In traditional theories, this phenomenon (referred to as integrative orientation, see e.g. Gardner 1985), is considered the main and most successful motivator for

foreign language learning. In further theoretical developments, Richard Clément defined the type and frequency of contact with native speakers of the L2 community as a prerequisite for the development of linguistic self-confidence, which in turn is assumed to affect motivation and language learning success (cf. Clément 1980, Clément & Kruidenier 1985, Sampasivam & Clément 2014).

In the course of globalisation, this view has been challenged especially with regard to English as a global lingua franca that has become rather detached from a particular speech community. In their studies in the Hungarian context, Dörnyei and colleagues provide evidence for a manifestation of the integrative orientation even in a context where L2 English learners have no contact at all with native speakers of English (Dörnyei et al. 2006). What remains unclear is the generalisability of these findings, i.e. if they apply to other contexts with different, less prestigious target languages.

3 Foreign language learning and L2 motivation research in Switzerland and beyond

Due to its multilingual reality, foreign language learning has a strong tradition in Switzerland. Analogous to developments in many European countries, an early introduction of foreign languages in school has become subject to controversial political discussions and the focus has changed from a rather local to an international perspective on foreign language learning curricula. Hence, English as a global lingua franca gained in relevance compared to the local official languages (see also Chapter 2, §2).

Several studies were conducted in order to inform and evaluate the implementation of the new foreign language learning curricula, and L2 motivation was one of the central objects of inquiry. The main findings suggest that primary school students are generally motivated to learn foreign languages and the ability to communicate with L2 speakers proved to be the strongest motivator (Bader & Schaer 2005, Husfeldt & Bader Lehmann 2009: 17–18, von Ow et al. 2012: 53, Kreis et al. 2014, Wiedenkiller & Lenz 2019: 56–57). At the same time, large-scale studies consistently indicate that English as the global lingua franca is generally favoured over French as a local foreign language (Stöckli 2004, Heinzmann 2010, 2013, Peyer et al. 2016, Brühwiler & Le Pape Racine 2017, Pfenninger & Singleton 2017). With respect to different motivational subcomponents, the results are more mixed: Whereas Heinzmann (2010: 14–15) and Peyer et al. (2016: 6–7) find higher values for English in all motivation components, Stöckli (2004: 59–60) did not identify differences between the two languages in extrinsic orientations,

Brühwiler & Le Pape Racine (2017: 173) even suggest that children are extrinsically more motivated for French than for English. As extrinsic or instrumental reasons for language learning appear to contribute less to L2 achievement (cf. Chapter 3, this volume), this result can be interpreted to the disadvantage of French, too.

The dominance of the global lingua franca over local languages is reflected in other contexts as well. For instance, Nikolov (2009: 103–104) shows that Hungarian learners of English are more motivated than learners of German (similar results are found in Csizér & Lukács 2010 for secondary school students). In a pan-European study, Busse (2017) suggests that – even at a young age – learners are well aware of the global status and prestige of English and that this can drastically affect attitudes towards the study of other foreign languages (cf. also discussion in Ushioda 2017). The complex interplay between English and local foreign languages and its impacts on early foreign language learning has further been discussed in Buyl & Housen (2014) in the Belgian context.

Motivational orientations also vary across individual and contextual factors. Firstly, gender effects have been documented for different contexts, with girls seeming generally more motivated than boys (Heinzmann 2009, Henry 2009, 2010, Brühwiler & Le Pape Racine 2017: 173–174; see also Dörnyei & Csizér 2002: 448, Courtney et al. 2017: 834) and gender differences are more pronounced in French than in English (Holder 2005, Brühwiler & Le Pape Racine 2017: 173–174; but see also Dewaele et al. 2016). Secondly, the linguistic repertoire is often discussed as a factor that shapes motivation in that multilingual children seem to be more motivated than monolinguals (e.g. Brühwiler & Le Pape Racine 2017: 174; some studies only find this effect for French, e.g. Stöckli 2004: 60, Heinzmann 2010: 19).

Beside these individual factors, the social learning environment plays a vital role in generating and maintaining motivation. Different findings suggest that teachers, peers, and parents are influencing motivation with their attitudes and support (Noels et al. 1999, Csizér & Kormos 2008, 2009, Husfeldt & Bader Lehmann 2009, Taguchi et al. 2009, Iwaniec 2014: 73, Peyer et al. 2016: 20–21, Pfenninger & Singleton 2016: 325–326, Busse 2017: 574, Sugita McEown et al. 2017, Wiedenkiller & Lenz 2019). Pfenninger & Singleton (2016: 336) conclude that it is just these collective daily experiences and relations in the language learning classroom that influence young learners L2 motivation substantially and that they should therefore play a crucial role in empirical investigations.

These general research trends are subject to some limitations. Due to theoretical and methodological discrepancies, findings from the above-mentioned studies are often difficult to compare. Above all, this concerns the construct of motivation in itself. Presumably clear-cut terms like “extrinsic”, “intrinsic”, “instrumen-

tal”, etc., are operationalised differently based on the authors’ theoretical viewpoint and this complicates the comparison of empirical findings. Furthermore, most studies focus on English, whereas less prestigious local foreign languages are often neglected¹. Finally, most of the existing data on young learners’ L2 motivation originate from areas far off language borders, regions where different language communities live closer together are under-researched, and supra-regional comparisons which allow for the investigation of contextual influences are lacking.

4 L2 motivation in the LAPS project

Aiming at bridging the research gaps discussed above, the present study explores the motivational dispositions of primary school students with similar school systems and language learning approaches but who differ in proximity to the local L2/L3 speech community.

By including a group of students who live close to the French-German language border as well as a group from a distant region, the relevance of the proximity to an L2 community can be examined.

In accordance with the description of curricular modalities (cf. Chapter 2, §2), the groups are henceforth referred to as group PROX (i.e. LAPS I – German-speaking children who live close to the French-German language border) and group DIST (i.e. LAPS II – children from a German-speaking region located far off the language border), respectively.

The analysis is led by the following research questions and hypotheses:

RQ1: How motivated are children to learn French and English at school with respect to different motivational subcomponents?

H1.1: Students are generally motivated to learn foreign languages at school. Based on theoretical considerations and previous studies, the motivation to use the L2 as a lingua franca is expected to be highest.

H1.2: According to previous findings, English is hypothesised to be clearly favoured over French. This disparity is expected to occur in all but extrinsic school-related motivations, because these goals (e.g. getting good grades or social recognition at school) are not expected to depend on the target language.

¹A special issue of the *Modern Language Journal* published in autumn 2017 shows that this is part of a wider debate and there is a growing awareness amongst scholars for the disparity between research on motivation for English and other languages (Ushioda & Dörnyei 2017).

RQ2: How do regional conditions affect students' L2 motivation when gender, language profile, teachers, class, and parents are controlled for?

H2: According to theories that attribute a central role to the proximity to the L2 community, an interaction between target language and student group is hypothesised: Whereas English is a "distant" language for both groups, they differ in geographical proximity to the French-speaking community. Thus, motivation for English is assumed to be approximately equal, but children close to the language border are assumed to be more motivated to learn French than those who live in a far-off region.

4.1 Method

4.1.1 Context, participants and design

To investigate motivational dispositions, data from the pupil questionnaire collected in spring 2018 and spring 2019 were analysed.

A total of 737 datasets from 5th and 6th graders of group PROX (LAPS I, T2, $n = 172$) and group DIST (LAPS II, T3, $n = 565$) were included in the analysis. Detailed information on the sample can be consulted in Chapter 2, §4. Sample sizes in this chapter deviate from the overall project, because only children who filled out the respective motivation questionnaires were considered for analysis. In addition, French and/or English native speakers' data were excluded from analysis of the respective L2 motivation constructs.

4.1.2 Pupil questionnaire

In order to measure affective dispositions, two pupil questionnaires with identical items for French and English were developed. According to theoretical considerations and with the target group of young learners in mind, the focus was set on different forms of extrinsic and intrinsic motivation, and in addition, children's ideal L2 self.²

Furthermore, as is argued e.g. by Ushioda (2009), individuals play an active role in shaping their own environments. Therefore, questions on teacher and

²Today, the majority of studies investigating affective dispositions rely on Dörnyei's L2MSS. At the same time, future self-representations as theorised in the L2MSS are assumed to be unstable at an early age and Dörnyei (2009) recommends the application of this framework from secondary school on. Therefore, the assessment of these constructs was not at the questionnaire's centre. However, three items on the ideal L2 self were integrated in order to test to what degree primary school children can already evaluate their future aspirations.

parental support were integrated in order to gauge students' perception and co-construction of their social context rather than modelling external influences as rigid background variables.

The scales were constructed based on previously validated questionnaires (Horwitz et al. 1986, Stöckli 2004, Dörnyei 2010, Heinzmann 2013, Peyer et al. 2016). Items were formulated as positive statements and were thematically grouped. Students were instructed to indicate their agreement with the presented statements on a 4-point Likert scale (1 = maximum disagreement, 4 = maximum agreement).

Cronbach's α were computed for all scales except for intrinsic motivation, which consists of two items only. The internal consistency was satisfactory for all scales (Cronbach's α ranges between 0.68 and 0.90, detailed information on internal consistency and a complete list of questionnaire items can be consulted in the analysis report, cf. <https://osf.io/gftvx/>).

4.1.3 Data analysis

Data preparation and analyses were conducted in R (R Core R Core Team 2019). The outcome variable was modelled as students' L2 motivation, comprising the multiple measurements per student of the five motivational subdimensions for the two target languages English and French. In order to deal with these multiple measurements per student and to view the students in relation to their context, mixed effects modelling was applied (e.g., Baayen et al. 2008, Pfenninger & Singleton 2016). The analysis report, datasets, and figures can be downloaded from <https://osf.io/fdr9v/>.

The package lme4 (Bates et al. 2015) was used to perform a linear mixed effects analysis of the relationship between motivational subdimensions and different target languages and learner groups. Model assumptions were tested through visual inspection. In order to assess the relevance of interactions and class effects, a complete model comprising all hypothesised main effects and interactions was fitted and compared to simpler models without the supposed interactions via likelihood-ratio tests.

5 Results and discussion

The final model contains the fixed effects and random intercepts displayed in Table 1 (models with random slopes did not converge). For categorical variables, the first mentioned represents the reference level.

Table 1: Fixed and random effects

Fixed effects	
Motivational dimension	Ideal L2 self, intrinsic motivation, extrinsic motivation: leisure motivation, school motivation, lingua franca motivation
Target language Group	English, French Contains information on region, school year, and number of English/French lessons at school – PROX5, PROX6, DIST5, DIST6
Multilingualism	Multilinguals (German and/or other L1), Monolinguals (German)
Gender	Boy, Girl
Teacher	Perceived teacher encouragement (centred, i.e. mean = 0)
Parents	Perceived parental encouragement (centred, i.e. mean = 0)
Random intercepts	
Subject	$N = 732$
Class	$N = 43$

The results of the final model are presented in Table 2. The intercept indicates the predicted mean on a scale of 1 to 4 for the respective reference levels displayed in Table 1 (Ideal L2 self, target language English, group PROX5, multilingual, boy). Hence, the predicted mean for the Ideal English self of a multilingual boy in group PROX5 is 3.31 on a scale of 1 to 4. Estimates for categorical fixed effects indicate changes with respect to the respective reference level. For example, the predicted outcome for intrinsic motivation is 0.63 lower than for the Ideal L2 self if all other parameters remain constant. Apart from main effects, an interaction was identified between target language and motivational subdimension, the results of which are indicated at the bottom of Table 2. In contrast, the analysis revealed no interaction between target language and group, indicating that there is no effect of the proximity to the target language community on motivation to learn French.

Effect plots for each fixed effect and the interaction between motivational dimension and target language are presented in Figures 1 to 4 and discussed according to the two research questions in §5.1 (L2 motivation for French and English with regard to motivational subdimensions) and §5.2 (effects of regional conditions on L2 motivation).

Table 2: Coefficients of mixed effects model Motivation ~ Dimension*Language + Multilingualism + Gender + Teacher + Parents + (1|Student) + (1|Class)

Fixed Effects	Est.	SE	t	CI 2.5%	97.50%	Random effects	Var	SD
(Intercept)	3.31	0.05	73.51	3.22	3.40	Student (Intercept)	0.09	0.30
Dimension Intrinsic	-0.63	0.04	-17.63	-0.69	-0.56	Class (Intercept)	0.008	0.09
Dimension Leisure	-1.00	0.04	-28.10	-1.07	-0.93	Residual	0.43	0.66
Dimension Lingua Franca	-0.05	0.04	-1.37	-0.12	0.02			
Dimension School	-0.77	0.04	-21.72	-0.84	-0.70			
Language French	-0.61	0.04	-17.18	-0.68	-0.54			
Group PROX5	-0.04	0.04	-0.98	-0.13	0.04			
Group PROX6	0.01	0.07	0.13	-0.14	0.16			
Multilingual yes	0.05	0.06	0.82	-0.07	0.17			
Gender girl	0.11	0.03	3.54	0.05	0.16			
Teacher	0.04	0.03	1.58	-0.01	0.10			
Parents	0.26	0.01	18.44	0.23	0.28			
	0.13	0.02	8.72	0.10	0.16			
Interaction Dimension:Language								
Intrinsic:French	0.22	0.05	4.40	0.12	0.32			
Leisure:French	-0.02	0.05	-0.46	-0.12	0.08			
LinguaFranca:French	-0.06	0.05	-1.17	-0.16	0.04			
School:French	0.74	0.05	14.84	0.64	0.84			

5.1 Research question 1: Young learners' L2 motivation for French and English with regard to motivational subdimensions

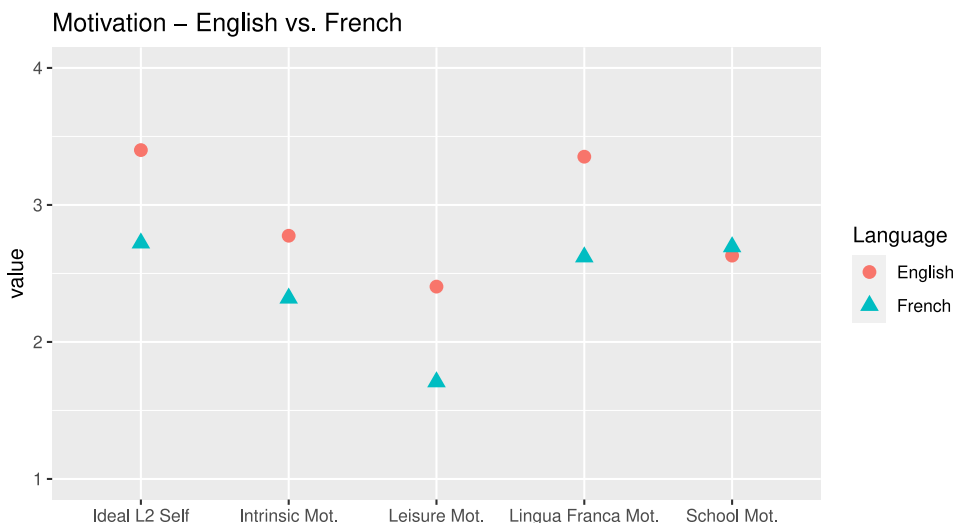


Figure 1: L2 Motivation according to target language and subdimension based on model predictions

In relation to the first research question, the predicted means for English and French learning motivation in the five subdimensions are presented in Figure 1. Our results are mostly in line with the hypotheses outlined in §4 and suggest that primary school students in the LAPS project are generally motivated to learn foreign languages at school, however, there are marked differences between target languages and between motivational subcomponents. As Figure 1 indicates, the pupils seem to be particularly motivated with regard to English, where all predicted means are around or above the scale mean. The ideal L2 self and the use of English for communicative purposes turn out as the two main motivators. In return, intrinsic motivation, as well as school and leisure motivation are lower.

With regard to French, the values are generally lower. Only school motivation is approximately at the same level as in English, all other values are substantially lower. This is how the interaction between motivational subdimension and target language can be interpreted: while school motivation is similar for both languages, the gap widens markedly in the other four dimensions.

Our results corroborate previous findings in that English learning motivation is generally higher than French learning motivation (Stöckli 2004, Heinzmann 2010, 2013, Peyer et al. 2016, Brühwiler & Le Pape Racine 2017, Pfenninger &

Singleton 2017). Based on Busse (2017), this difference suggests that even young learners consciously deal with the relation of these two languages and that they assign a more valuable status to English than to French (cf. also Ushioda 2017, Buyl & Housen 2014). As one of the anonymous reviewers pointed out, a further explanation for the preference for English might lay in the (perceived) difficulty of the target language, in that German-speaking students rate English as easier to learn than French.

Irrespective of the gap between the target languages, our results encouragingly support previous studies with regard to the use of the target language as a *lingua franca*, which seems to be one of the students' main aspirations for foreign language learning (cf. e.g. Schaer & Bader 2003, Stöckli 2004). Likewise, students reported high values for the ideal L2 self, which has rarely been studied with young learners (see Henry 2009 for an exception, cf. §3 in this Chapter and §2, Chapter 8). This suggests that students have vivid future self-guides and that they can imagine using English and French competently in the future. In addition, the Cronbach's α of these scales (0.81, 0.87) indicate a consistent response behaviour. However, viewing this as an entirely positive finding would be premature, because young students have been found to overestimate themselves and Dörnyei (2009: 38) himself argues that learners' future self representations are not stable at this early age.

Similarity in school-related motivation in both languages has also been found in previous studies (e.g. Stöckli 2004). This is not surprising, insofar as it is related to an instrumentalisation of the languages with the goal of academic success. Good educational performance proves to be important for the students in our sample, irrespective of the target language.

Furthermore, low values for leisure motivation are consistent with previous findings (cf. e.g. Schaer & Bader 2003, Stöckli 2004, Heinzmann 2013) and this result is not surprising. English was regarded as much more important than French, which can be linked to the greater benefits in terms of its use on the internet and in computer games. At the same time, the comparatively weaker Cronbach's α and the fact, that even the values for English are comparatively low, suggest that such goals might not yet be highly relevant for primary school students.

A point that requires further discussion is intrinsic motivation. In accordance with other studies in similar contexts, intrinsic motivation for English is higher than for French. But in our data, intrinsic motivation for English is only slightly above the scale mean, the values for French are slightly below. Based on Deci & Ryan's (1985) SDT and on the repeatedly confirmed influence on language learning success, this is a rather discouraging result.

5.2 Research question 2: The closer, the better?

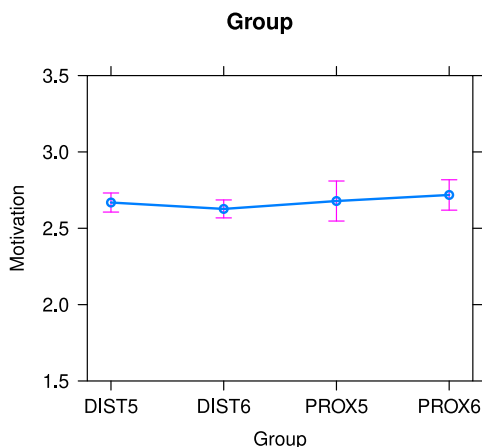


Figure 2: Motivation according to learner group

With regard to the second research question, the model reveals no main effect for regional conditions and no interaction between regional conditions and target language. As Figure 2 as well as the estimates displayed in Table 2 suggest, all learner groups are equally motivated, irrespective of their distance to the L2 community.

According to theories that attribute a central role to the proximity to the L2 community (such as Gardner & Lambert’s socio-educational model or Clément’s contact hypothesis, see Section 2 above), an interaction between target language and student group was hypothesised: Whereas English is a “distant” language for both groups, they differ in geographical proximity to the French-speaking community. Thus, motivation for English would be assumed to turn out approximately the same, whereas motivation for French would be expected to be higher for children close to the French language border compared to their peers living in a region far off.

Explanations for the absence of this interaction, and thus, the absence of an effect of proximity to a speech community on the learners’ L2 motivation could be linked to the global status of the two languages being more important than the actual geographical proximity. As already mentioned above, the criticism of theories by Gardner and his associates on the relevance of a particular L2 speech community might be transferable from English to other languages (cf. also discussions in Busse 2017 and Ushioda 2017).

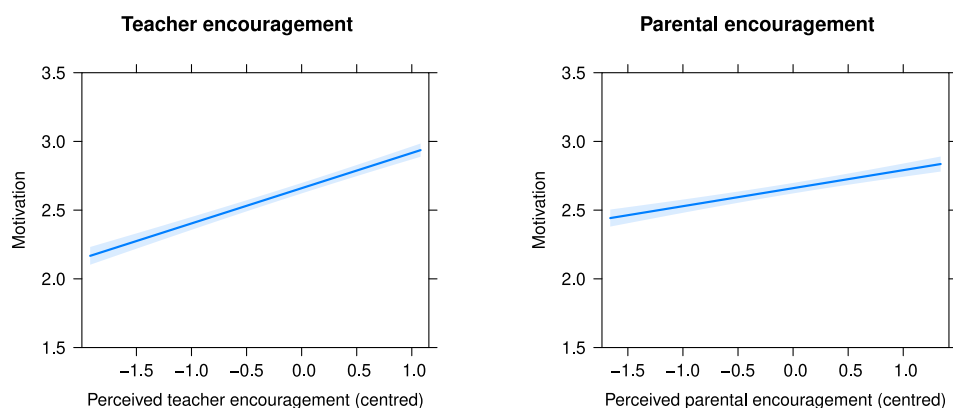


Figure 3: Teacher and parental encouragement effect plots

Concerning the control variables entered in the model, perceived support by the L2 teacher seems to have a considerable impact on students' L2 motivation (cf. Figure 3, left plot). To a lesser extent, this also applies to parental encouragement (cf. Figure 3, right plot). However, based on our data, no conclusion on the causal direction of these effects can be drawn (e.g. children who were more motivated in the first place might also perceive parental and teacher encouragement more positively).

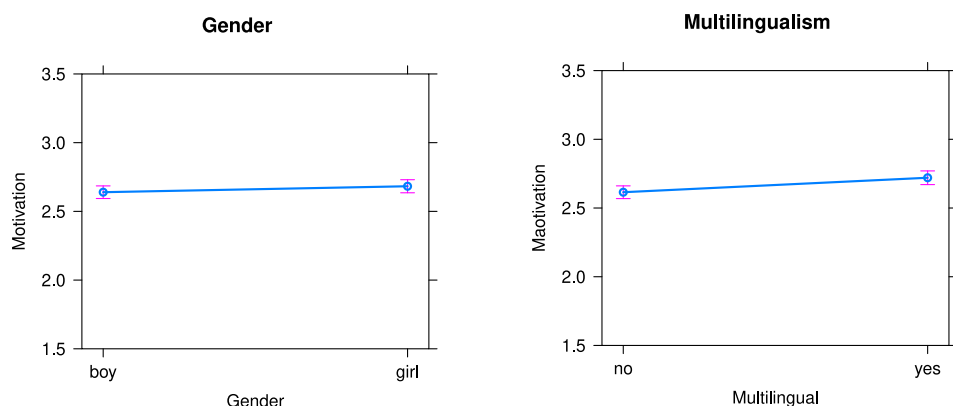


Figure 4: Multilingualism and gender effect plots

In contrast, effects of multilingual background or gender were marginal (cf. Figure 4), and no notable between-class variation was revealed (cf. Table 2, random effect "Class").

Findings from Pfenninger & Singleton (2016) give reason to assume that the educational environment (teachers, their methodologies, peers, etc.) exerts a great influence on the individual pupils in the classroom and, thus, differences between classes should become apparent. It is therefore surprising that no evidence for substantial between-class variation was found in the current study. However, this does not imply that the teacher has no impact on his or her students. From the control variables in our model, perceived motivation and support by the teacher is the one with the greatest impact on students' motivation. The fact that the impact of parental encouragement is comparatively lower is in line with previous findings (e.g. Holder 2005), who showed in a qualitative approach that parental influence becomes weaker than the school environment for children at the end of primary school.

In relation to multilingualism or gender, our data does not lend support to previous findings (Dörnyei & Csizér 2002, Holder 2005, Heinzmann 2009, 2010, Brühwiler & Le Pape Racine 2017, Henry 2009, Courtney et al. 2017). Whereas numerically, girls and children with a multilingual background are more motivated, these effects are modest and uncertain.

6 Limitations

The results of this study are subject to several limitations. Firstly, intrinsic motivation, although representing an important construct, was assessed by two items only. In order to get a more detailed view, it would be useful to gauge this construct more extensively in future studies. Secondly, the distance to the L2 community was solely based on geographical distance, direct contact with speakers of the respective community was not integrated in the questionnaire. An assessment of the intensity of interactions and contact with L2 speakers could possibly clarify the extent to which pupils actually seize opportunities to get in touch with L2 speakers and if this has an impact on their L2 motivation. Finally, the quantitative approach taken in this chapter along with a big sample ($n=737$) enables for generalising conclusions. At the same time, it is obvious, that such a wide-ranging construct as language learning motivation cannot be assessed comprehensively via quantitative questionnaire items alone. Additional qualitative research (e.g. via semi-structured interviews with specific learners or learner groups) could certainly provide further valuable insights.

7 Conclusion

With the limitations discussed above considered, a series of conclusions can be drawn from this study. In general, our data suggests that even in an educational environment, where foreign languages are part of the mandatory curriculum, pupils have various motivations for language learning beyond academic success. For instance, future self-representations and the active use of the languages for communication play a crucial role as motivators in the L2 learning process. This particularly applies to English, which is strongly favoured over French in all but the educational subdimension. The most striking result is related to the distance to an L2 community. Contrary to our expectations, the results of the current study don't support the conclusion that proximity to the French speaking community fosters motivation to learn the respective language. As already discussed above, this might possibly be explained by the mandatory school context and the children attributing a higher status to English than to French. A closer look at the actual willingness to and intensity of contact to L2 speakers could shed more light on this intriguing finding. As to the other factors integrated in the analysis, the teacher proved to be the most important social agent in fostering pupils' L2 motivation, whereas parents seem to exert less influence. Finally, no considerable effect of multilingualism and gender could be identified in the current study.

References

- Baayen, R. Harald, Doug J. Davidson & Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4). 390–412.
- Bader, Ursula & Ursula Schaer. 2005. *Evaluation Englisch in den 6. Klassen Appenzell Innerrhoden 2005*. Aarau: PH FHNW.
- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: 10.18637/jss.v067.i01.
- Brühwiler, Christian & Christine Le Pape Racine. 2017. Entwicklung motivationaler Orientierungen beim multiplen Erwerb von Französisch und Englisch als Fremdsprachen am Übergang von der Primarstufe in die Sekundarstufe I. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 22(1). 167–181.
- Busse, Vera. 2017. Plurilingualism in Europe: Exploring attitudes toward English and other European languages among adolescents in Bulgaria, Germany, the Netherlands, and Spain. *The Modern Language Journal* 101(3). 566–582. DOI: 10.1111/modl.12415.

- Buyl, Aafke & Alex Housen. 2014. Factors, processes and outcomes of early immersion education in the francophone community in Belgium. *International Journal of Bilingual Education and Bilingualism* 17(2). 178–196. DOI: 10.1080/13670050.2013.866627.
- Carroll, John B. 1958. A factor analysis of two foreign language aptitude batteries. *The Journal of General Psychology* 58(1). 3–19. DOI: 10.1080/00221309.1958.9710168.
- Carroll, John B. 1964. The prediction of success in intensive foreign language training. In *Training research and education*, 87–136. Pittsburgh: University of Pittsburgh Press.
- Clément, Richard. 1980. Ethnicity, contact and communicative competence in a second language. In Howard Giles, Peter W. Robinson & Philip M. Smith (eds.), *Language*, 147–154. Amsterdam: Pergamon. DOI: 10.1016/B978-0-08-024696-3.50027-2.
- Clément, Richard & Bastian G. Kruidenier. 1985. Aptitude, attitude and motivation in second language proficiency: a test of Clément's model. *Journal of Language and Social Psychology* 4(1). 21–37.
- Courtney, Louise, Suzanne Graham, Alan Tonkyn & Theodoros Marinis. 2017. Individual differences in early language learning: a study of English learners of French. *Applied Linguistics* 38(6). 824–847. DOI: 10.1093/applin/amv071.
- Csizér, Kata & Judit Kormos. 2008. The role of intercultural contact in learning German in Hungary: a structural equation modelling approach. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 13(2). 14 S.
- Csizér, Kata & Judit Kormos. 2009. Modelling the role of inter-cultural contact in the motivation of learning English as a foreign language. *Applied Linguistics* 30(2). 166–185. DOI: 10.1093/applin/amn025.
- Csizér, Kata & Gabriella Lukács. 2010. The comparative analysis of motivation, attitudes and selves: the case of English and German in Hungary. *System* 38(1). 1–13. DOI: 10.1016/j.system.2009.12.001.
- Deci, Edward L. & Richard M. Ryan. 1985. *Intrinsic motivation and self-determination in human behaviour*. New York: Plenum.
- Dewaele, Jean-Marc, Peter Macintyre, Carmen Boudreau & Livia Dewaele. 2016. Do girls have all the fun? anxiety and enjoyment in the foreign language classroom. *Theory and Practice of Second Language Acquisition* 2. 41–63.
- Dörnyei, Zoltán. 2005. Motivation and “self-motivation”. In Zoltán Dörnyei (ed.), *The psychology of the language learner individual differences in second language acquisition*, 65–118. Mahwah, NJ: Lawrence Erlbaum Associates.

- Dörnyei, Zoltán. 2009. The L2 motivational self system. In Zoltán Dörnyei & Ema Ushioda (eds.), *Motivation, language identity and the L2 self*. Bristol: Multilingual Matters.
- Dörnyei, Zoltán. 2010. *Questionnaires in second language research*. 2nd edn. New York: Routledge.
- Dörnyei, Zoltán & Kata Csizér. 2002. Some dynamics of language attitudes and motivation: results of a longitudinal nationwide survey. *Applied Linguistics* 23(4). 421–462.
- Dörnyei, Zoltán, Kata Csizér & Nora Németh. 2006. *Motivation, language attitudes and globalisation: a hungarian perspective*. Clevedon: Multilingual Matters.
- Gardner, Robert C. 1985. *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.
- Gardner, Robert C. & Wallace E. Lambert. 1959. Motivational variables in second language acquisition. *Canadian Journal of Psychology* 13). 266–272.
- Heinzmann, Sybille. 2009. "girls are better at language learning than boys": do stereotypic beliefs about language learning contribute to girls' higher motivation to learn english in primary school? *Bulletin VALS-ASLA*.
- Heinzmann, Sybille. 2010. Hat die einföhrung von "fröhenglisch" in der primarschule einen einfluss auf die motivation der primarschulkinder französisch zu lernen? *Bulletin VALS-ASLA* 91. 7–27.
- Heinzmann, Sybille. 2013. *Young language learners' motivation and attitudes: Longitudinal, comparative and explanatory perspectives*. London: Bloomsbury.
- Henry, Alastair. 2009. Gender differences in compulsory school pupils' L2 self-concepts: A longitudinal study. *System* 37(2). 177–193.
- Henry, Alastair. 2010. Contexts of possibility in simultaneous language learning: Using the L2 motivational self system to assess the impact of global English. *Journal of Multilingual and Multicultural Development* 31(2). 149–162. DOI: 10.1080/01434630903471439.
- Holder, Martin C. 2005. *Fähigkeitskonzept und leistungsmotivation im fremdsprachenunterricht*. Bern: Peter Lang.
- Horwitz, Elaine, Michael B. Horwitz & Joann Cope. 1986. Foreign language classroom anxiety. *The Modern Language Journal* 70(2). 125–132.
- Husfeldt, Vera & Ursula Bader Lehmann. 2009. *Englisch an der primarschule. lernstandserhebung im kanton aargau*. Aarau: BKS.
- Iwaniec, Janina. 2014. Motivation of pupils from southern poland to learn english. *System* 45. 67–78. DOI: 10.1016/j.system.2014.05.003.

- Kreis, Annelies, Michaela Williner & Christoph Maeder. 2014. *Englischunterricht in der Primarschule des Kantons Thurgau: Schlussbericht der Evaluation*. Kreuzlingen: Pädagogische Hochschule Thurgau.
- Nikolov, Marianne. 2009. Early modern foreign language programmes and outcomes: factors contributing to hungarian learners' proficiency. In Marianne Nikolov (ed.), *Early learning of modern foreign languages. processes and outcomes*, 90–107. Bristol, Buffalo, Toronto: Multilingual Matters.
- Noels, Kimberly A. 2001. New orientations in language learning motivation: towards a model of intrinsic, extrinsic, and integrative orientations and motivations. In Zoltán Dörnyei & Richard Schmidt (eds.), *Motivation and second language acquisition*, 43–68. Honolulu: University of Hawaii Press.
- Noels, Kimberly A., Richard Clément & Luc G. Pelletier. 1999. Perceptions of teachers' communicative style and students' intrinsic and extrinsic motivation. *The Modern Language Journal* 83(1). 23–34. DOI: 10.1111/0026-7902.00003.
- Peyer, Elisabeth, Mirjam Andexlinger, Karolina Kofler & Peter Lenz. 2016. *Projekt Fremdsprachenevaluation BKZ: Schlussbericht zu den Sprachkompetenztests*. Tech. rep. Université de Fribourg. <http://doc.rero.ch/record/259287/> (9 February, 2018).
- Pfenninger, Simone & David Singleton. 2016. Affect trumps age: a person-in-context relational view of age and motivation in SLA. *Second Language Research* 32(3). 311–345. DOI: 10.1177/0267658315624476.
- Pfenninger, Simone & David Singleton. 2017. *Beyond age effects in instructional L2 learning: Revisiting the age factor*. Bristol: Multilingual Matters.
- R Core Team. 2019. *R: A language and environment for statistical computing*. Version 3.5.1. Wien, Österreich. <https://www.R-project.org/>.
- Ryan, Richard M. & Edward L. Deci. 2002. An overview of self-determination theory: an orgasmic dialectal perspective. In Edward L. Deci & Richard M. Ryan (eds.), *Handbook of self-determination research*, 3–33. New York: University of Rochester Press.
- Sampasivam, Sinthujaa & Richard Clément. 2014. The dynamics of second language confidence: contact and interaction. In Sarah Mercer & Marion Williams (eds.), *Multiple perspectives on the self in SLA*, 23–40. Bristol, Buffalo, Toronto: Multilingual Matters.
- Schaer, Ursula & Ursula Bader. 2003. *Evaluation Englisch an der Primarschule. Projekt 012. Appenzell Innerrhoden. Mai bis November 2003*. Zürich, Luzern: PHZH, PHZ Luzern, ILLB Bern.
- Stöckli, Georg. 2004. *Motivation im Fremdsprachenunterricht: Eine theoriegeleitete empirische Untersuchung in 5. und 6. Primarschulklassen mit Unterricht in*

- Englisch und Französisch* (Pädagogik bei Sauerländer). Oberentfelden: Sauerländer Verlag.
- Sugita McEown, Maya, Yasuyo Sawaki & Tetsuo Harada. 2017. Foreign language learning motivation in the Japanese context: social and political influences on self. *The Modern Language Journal* 101(3). 533–547. DOI: 10.1111/modl.12411.
- Taguchi, Tatsuya, Michael Magid & Mostafa Papi. 2009. The L2 motivational self system among Japanese, Chinese and Iranian learners of English: A comparative study. In Zoltán Dörnyei & Ema Ushioda (eds.), *Motivation, language identity and the L2 self*, 66–97. Bristol, Buffalo, Toronto: Multilingual Matters.
- Ushioda, Ema. 2009. A person-in-context relational view of emergent motivation, self and identity. In Zoltán Dörnyei & Ema Ushioda (eds.), *Motivation, language identity and the L2 self*, 215–228. Bristol, Buffalo, Toronto: Multilingual Matters.
- Ushioda, Ema. 2017. The impact of global English on motivation to learn other languages: toward an ideal multilingual self. *The Modern Language Journal* 101(3). 469–482. DOI: 10.1111/modl.12413.
- Ushioda, Ema & Zoltán Dörnyei. 2017. Beyond global English: motivation to learn languages in a multicultural world: introduction to the special issue. *The Modern Language Journal* 101(3). 451–454. DOI: 10.1111/modl.12407.
- von Ow, Anna, Vera Husfeldt & Ursula Bader-Lehmann. 2012. Einflussfaktoren für den lernerfolg von englisch an der primarschule. *Babylonia* (1). 52–57.
- Wiedenkiller, Eva & Peter Lenz. 2019. *Schlussbericht zum Projekt "Ergebnisbezogene Evaluation des Französischunterrichts in der 6. Klasse (HarmoS 8) in den sechs Passepartout-Kantonen"; durchgeführt von Juni 2015 bis März 2019 am Institut für Mehrsprachigkeit der Universität und der Pädagogischen Hochschule Freiburg im Auftrag der Passepartout-Kantone*. Schweizerischer Dokumentenserver Bildung. <http://edudoc.ch/record/206193> (12 March, 2020).

Chapter 8

The dynamics of young learners' L2 motivation: A longitudinal perspective

Carina Steiner

University of Bern, Center for the Study of Language and Society

Affective dispositions in language learning are regarded as dynamic constructs that change over time. Yet, longitudinal studies in this field are relatively rare. In order to better understand these changes, this chapter reports a study that investigates the development of 578 primary school students' English and French learning motivation, self-concepts, anxiety, and perceived support of teachers and parents at three measurement times over two academic years. While no drastic overall changes could be identified in our data, affective dispositions remain constantly higher for English than for French. Whereas students maintain their strong English-related motivations and self-concepts and English learning anxiety remains on a low level, French-related intrinsic motivation and self-concepts weaken with a parallel growth of anxiety. With regard to extrinsic motivation, both in leisure and school contexts, as well as perceived teacher and parental encouragement, a slight decrease is observed in both target languages.

These results are discussed with regard to previous findings and in light of the relation between English as a global lingua franca and French as a local foreign language.

1 Introduction, theoretical remarks

This chapter deals with a more detailed examination of the *dynamics* of different motivational and related affective dispositions in language learning at the primary school level. More precisely, the focus is set on the long-term dynamics of intrinsic motivation, motivation to use the L2 as a lingua franca, leisure- and school-related extrinsic forms of motivation, the ideal L2 self, as well as current self-concepts, perceived teacher and parental encouragement, and foreign



Carina Steiner. 2021. The dynamics of young learners' L2 motivation: A longitudinal perspective. In Raphael Berthele & Isabelle Udry (eds.), *Individual differences in early instructed language learning: The role of language aptitude, cognition, and motivation*, 163–178. Berlin: Language Science Press.

DOI: 10.5281/zenodo.5464759 

language learning anxiety. General theoretical and empirical underpinnings on these constructs can be consulted in Chapters 1 and 7 in this volume.

Researchers in the field of L2 motivation commonly agree on the fact that these constructs change – at least to a certain extent – over time. For example, Gardner (2007: 11) suggests that, although not considered as personality traits, motivational dispositions are relatively stable, and yet “amenable to change under certain conditions”. Dörnyei & Ottó (1998: 65) go further and define motivation in a general sense as “dynamically changing cumulative arousal in a person”.

Related to the question whether this construct changes over time, the temporal perspective on motivation is crucial: While moment-by-moment experiences on the micro-level (e.g. motivational fluctuations during a specific task or task cycle) are prone to change, processes on the macro-level (e.g. general affective dispositions towards L2 learning or the L2 speech community) remain more stable over a longer period of time (Dörnyei & Ushioda 2011: 6). This chapter aims to provide a long-term view on L2 motivation development and, thus, all subsequent discussions relate to motivational dispositions on the *macro*-level which are less fluctuant than situational motivations on the micro-level.

Even though the importance of various affective dispositions for the language learning process has been widely discussed, very few studies deal with their dynamics at an early age. This is due to the assumption that motivation, attitudes, and other affective dispositions are only emerging at a certain age and are therefore instable in younger learners (see e.g. Gardner 2006, Dörnyei 2009). As Heinzmann (2013) argues from a pedagogical perspective, this is exactly why more empirical research into young learners’ motivation and attitudes is needed:

[I]f young children’s motivation and attitudes are not yet stabilized, this means that they are still malleable and something can be done about negative motivational and attitudinal dispositions before they become firmly established. (Heinzmann 2013: 4)

2 Motivational dynamics at the primary school level

Although various scholars have pointed out the need for investigating the dynamics of children’s language learning motivation and related affective dispositions (see e.g. Cenoz 2004, McGroarty 2001), there are only few studies with a longitudinal design enabling robust conclusions about developmental processes (see e.g. Mihaljević Djigunović 2012 or Mihaljević Djigunović & Nikolov 2019 for an overview).

An insightful study for the Swiss context comes from Heinzmann (2013), who investigated primary school students' English learning motivation and attitudes at three measurement points in 3rd, 4th and 5th grade ($n = 552$). Her results suggest that pupils started with a strong motivation that remained relatively stable over the three measurement points. A closer examination of different subdimensions revealed that the biggest changes manifested in a slight drop of the intrinsic motivation from 3rd to 4th grade, while the motivation to use English as a lingua franca became continuously more important over time (cf. Heinzmann 2013). Similar results in relation to instrumental motivation, which overlaps with Heinzmann's lingua franca motivation, were also found by Nikolov (2002) or, in a cross-sectional design, by Tragant (2006: 256).

Indices for an overall decline of motivation during primary school can be found in Bader & Schaer (2005; see also Schaer & Bader 2003, or Cenoz 2004 for similar tendencies for children in the Basque country). In contrast, Pfenninger & Singleton (2016: 324) show positive development tendencies for older students between 13 and 18 years of age.

In addition to motivational changes, Stöckli (2004: 93) investigated the English self-concept of 133 pupils in a longitudinal perspective. His results suggest a decline after the 1st grade, before re-strengthening until the end of 5th grade. Stöckli argues that young children tend to overestimate themselves at the beginning of the learning process, but growing experience reassures the learners continuously (for a related discussion, see also Mihaljević Djigunović & Lopriore 2011: 50).

A longitudinal study by Henry (2009) investigated young learners' future self-concepts in the sense of Dörnyei's L2 motivational self system (e.g. Dörnyei 2009; cf. Chapter 1, §4). In Henry's study, a motivation questionnaire was administered to Swedish students of English ($n = 169$) after one and four years of instruction. While the results suggest a stable development for the entire sample, further analyses revealed different trajectories according to gender: whereas girls' ideal self strengthened from 6th to 9th grade, it weakened for boys (cf. Henry 2009: 184).

As a final important factor, foreign language anxiety (cf. MacIntyre 1999) is negatively related to persistence in foreign language learning.

Similar to the study by Henry (2009), a large-scale cross-sectional study by Dewaele et al. (2016) revealed gender effects related to foreign language anxiety. The results revealed that, across all age groups included in the analysis, girls tended to experience not only more positive, but also more negative emotions in foreign language learning. The authors conclude that this generally higher emotional activation in both negative and positive ways leads to more promising learning outcomes for girls than for boys (cf. Dewaele et al. 2016: 59).

3 Motivational dynamics and the relation between English and French as foreign languages

The discrepancy between motivation to learn English vis-à-vis other foreign languages has been widely discussed in Switzerland and beyond (see e.g. Buyl & Housen 2014, Ushioda 2017, Busse 2017). Whereas the gap between English and French is discussed and supported by results in Chapter 7, the perspective in this chapter is developmental in nature. Scholars have repeatedly pointed out that pupils who have already learned English at school can lose their motivation to learn any other foreign language because, from an instrumental view, the global lingua franca serves well enough as a communication tool (see e.g. Hufeisen 2003: 9, Meissner et al. 2008: 109, Ushioda 2017).

In the Swiss context, this concern has not been supported by empirical findings. In a quasi-experimental study carried out in Central Switzerland, no negative effect of previous English instruction on French learning motivation could be detected (Heinzmann 2010). These results were supported by Brühwiler & Le Pape Racine (2017: 176) who investigated motivational changes in relation to both English and French at the transition from primary to secondary school.

With regard to the (in)stability of French- and English-related motivation, instrumental reasons referring to the aspiration for academic or professional success through language learning, were found to remain stable. Similar to Heinzmann (2013), a decline of intrinsic motivation was observed, with a parallel rise of motivation to use English as a lingua franca. The pattern was slightly different for French, where no growth of communication motivation could be identified (for a study with a similar age cohort in the UK, see Graham et al. 2016).

In relation to foreign language anxiety, insights can be drawn from a longitudinal study conducted within a monitoring programme of the new foreign language curricula in regions at the French-German language border in Switzerland (cf. Singh & Elmiger 2017). Questionnaire-based data collected between grade 5 and 8 revealed that, while children were continuously more satisfied with their English instruction, they felt much more stressed in French at the first data collection. Interestingly, stress levels tended to slightly lower for French and to increase for English over time. However, these results are to be interpreted cautiously because the sample changed between different measurement points and questionnaires were partly adapted.

4 Motivational dynamics in the LAPS project

As outlined above, there are few studies that investigate language learning motivation with a longitudinal design and even fewer with a focus on the relation between English and other, less prestigious foreign languages. As discussed in Chapter 7, Swiss children usually consider learning English more attractive than learning French. This leads us to define French as a less prestigious foreign language, at least in this particular context.

This chapter aims to provide additional insights by examining the dynamics of primary school students' motivation, self-concept, anxiety, as well as perceived teacher and parental encouragement with regard to English and French learning at school.

The following research question is addressed: *How do primary school students' motivation, as well as associated affective dispositions, change over a period of two academic years with regard to English and French?*

4.1 Participants and instruments

In order to investigate the dynamics of affective dispositions, a total of nine scales from the pupil questionnaire in the LAPS II project administered in autumn 2017 (T1), spring 2018 (T2), and spring 2019 (T3) are analysed. A total of 578 datasets are included in the analysis (detailed information on participants can be consulted in Chapter 2, §4, in this volume; the sample size deviates from the overall data due to the exclusion of native speakers of English and/or French and due to missing questionnaire data of some pupils).

At T1, students were at the beginning of grade 4 (cohort 4, $n = 273$) and grade 5 (cohort 5, $n = 305$). At T2, they were at the end of grade 4 and 5, and at T3 at the end of grade 5 and 6, respectively. Due to the integration of two age cohorts, the measurement points are partly overlapping in relative terms of school years. Table 1 displays this overlap.

4.2 Questionnaires

In the questionnaires related to affective dispositions, children were instructed to indicate for each item on a scale of 1 to 4 how strongly they agree with a certain statement.¹

¹1 = maximum disagreement; 4 = maximum agreement; see analysis report for a full description of all scales: <https://osf.io/r5ypz/>

Table 1: Overlap in measurement points between the two age cohorts in LAPSII

Cohort 4	T1: beginning of 4 th grade	T2: end of 4 th grade	T3: end of 5 th grade	
Cohort 5		T1: beginning of 5 th grade	T2: end of 5 th grade	T3: end of 6 th grade

The English Questionnaire was administered at all three data collections. However, some items were not repeated based on a factor analysis at T1. For the calculation of summary scores, only items that were asked at all three measurement points were integrated.

The French Questionnaire was administered at T2 and T3. Due to the fact that children in cohort 4 did not start French classes before T3, they were excluded from analyses related to French learning motivation.

Cronbach's α were computed for all scales except for intrinsic motivation, which consists of two items only. The internal consistency was satisfactory for all scales (Cronbach's α ranges between 0.67 and 0.91, cf. analysis report for detailed information, <https://osf.io/mc8jf/>).

4.3 Data analysis

The data were processed and analysed in R (R Core R Core Team 2019). Datasets and reports can be downloaded from OSF (<https://osf.io/mc8jf/>). Changes in affective dispositions over time were analysed in mixed effects models (package *lme4*, Bates et al. 2015) that account for multiple measurements per student and allow for the test of potential class effects. For each affective subdimension, a model was fitted with the respective summary score as dependent variable. Group (cohort 4 and 5) and measurement point (T1, T2, T3) were entered as fixed effects. Random intercepts were entered for students and classes. Model assumptions were visually tested. In order to examine the relevance of the effects and potential interactions, a comprehensive model containing all effects and interactions was calculated and then compared to different reduced models via likelihood ratio tests (cf. Winter 2013).

5 Results

5.1 L2 motivation, self-concepts, and anxiety

The results of all final models related to English are summarised in Table 2 (cf. analysis report for full output of each model: <https://osf.io/r5ypz/>). The intercept indicates the predicted mean for cohort 4 at the first measurement point on a scale of 1 to 4. Estimates for the fixed effect *group* indicate discrepancies between cohort 4 and 5 at T1, effects of *measurement point* (T2 and T3) signal changes in the respective affective dimension over time. An interaction between group and measurement point was only revealed for intrinsic motivation, indicating different developments for the two age cohorts over time (cf. Figure 1 and discussion below).

Between-subject variation and between-class variation is displayed under the respective random effect where applicable. The ideal L2 self was assessed at T2 and T3 only, and thus, the intercept indicates the predicted mean for T2.

Table 2: Fixed and random effects for English motivation, self-concepts, and anxiety. AD: Affective dispositions, IM: Intrinsic motivation, SM: School motivation, LM: Leisure motivation, LFM: Lingua Franca motivation, IS: Ideal L2 Self, SC: L2 self-concept, AN: Anxiety, Int.: Intercept.

AD	Int.	Fixed effects (\pm SE)			Random effects	
		Group Cohort 5	Measurement point T2	T3	School ^a SD	Class ^b SD
IM	2.78	-0.08 ± 0.07	0.02 ± 0.05	-0.11 ± 0.05	0.55	
SM	2.97	-0.23 ± 0.07	-0.13 ± 0.04	-0.24 ± 0.04	0.53	0.16
LM	2.65	-0.12 ± 0.06	-0.12 ± 0.04	-0.18 ± 0.04	0.57	
LFM	3.26	0.09 ± 0.06	0.02 ± 0.03	0.04 ± 0.03	0.39	0.12
IS	3.36	0.05 ± 0.06		0.02 ± 0.03	0.43	0.10
SC	2.98	0.08 ± 0.05	0.01 ± 0.03	0.03 ± 0.03	0.53	
AN	2.09	-0.13 ± 0.06	-0.03 ± 0.03	-0.02 ± 0.03	0.61	

^a*n* = 578

^b*n* = 32

Table 3 presents the results for changes in affective dispositions related to French for cohort 5 (as mentioned above, cohort 4 was excluded from these analyses because they did not attend French classes at T2). Estimates for *measurement*

point indicate changes in affective dispositions between T2 and T3. Between-subject and between-class variation is indicated under the respective random effects where applicable.

Table 3: Fixed and random effects for French motivation, self-concepts, and anxiety. AD: Affective dispositions, IM: Intrinsic motivation, SM: School motivation, LM: Leisure motivation, LFM: Lingua Franca motivation, IS: Ideal L2 Self, SC: Current self-concept, AN: Anxiety, Int.: Intercept.

AD ^a	Int.	Fixed effects	Random effects	
		Measurement point	Student ^b	Class ^c
		T3 (\pm SE)	SD	SD
IM	2.43	-0.26 ± 0.05	0.67	
SM	2.75	-0.14 ± 0.05	0.60	
LM	1.69	-0.09 ± 0.05	0.45	
LFM	2.69	-0.17 ± 0.05	0.53	
IS	2.78	-0.27 ± 0.04	0.57	
SC	2.53	-0.16 ± 0.04	0.56	
AN	1.97	0.13 ± 0.05	0.59	

^aCohort 5 only.

^b $n = 305$

^c $n = 19$

All effects based on model predictions for intrinsic, school, leisure, and lingua franca motivation as well as for the ideal L2 self, the current L2 self-concept and anxiety are plotted in Figure 1 (English) and Figure 2 (French).

As the data in Table 2 and Figure 1 suggest, affective dispositions related to English remain relatively stable from the beginning of grade 4 until the end of primary school. The pupils start with a particularly strong lingua franca motivation and a high English self-concept, and these values even marginally increase over time. The Ideal L2 Self, which was measured at T2 and T3, also starts out very high and remains stable over the course of a school year. English learning anxiety, in contrast, remains on a low level, even decreasing slightly over time. In contrast, extrinsic motivations related to school and leisure activities are less stable, they weaken for both age cohorts from T1 to T3.

The only dimension where an interaction between group and measurement point could be identified is the intrinsic motivation: While this motivation

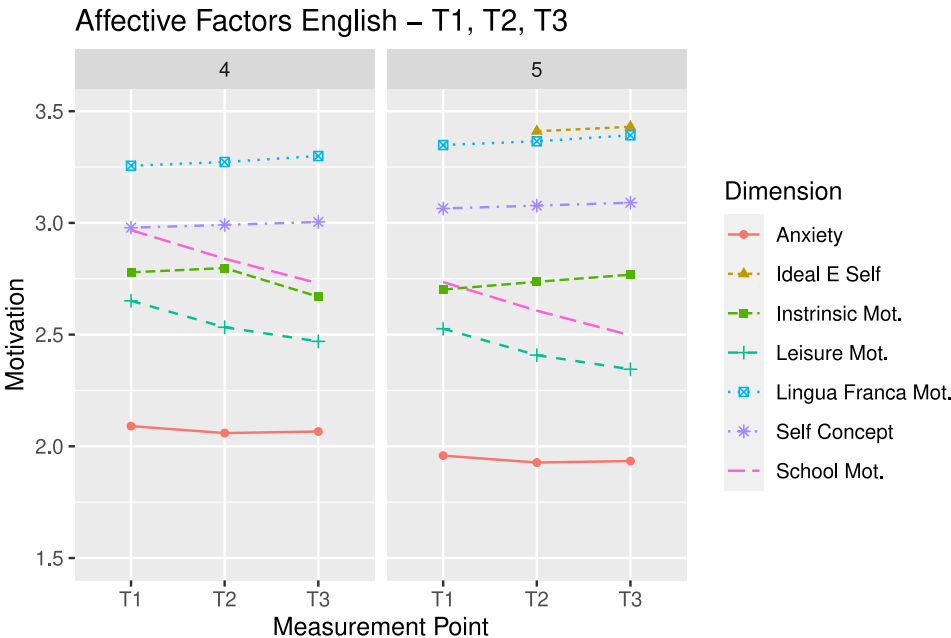


Figure 1: Dynamics of affective dispositions related to English

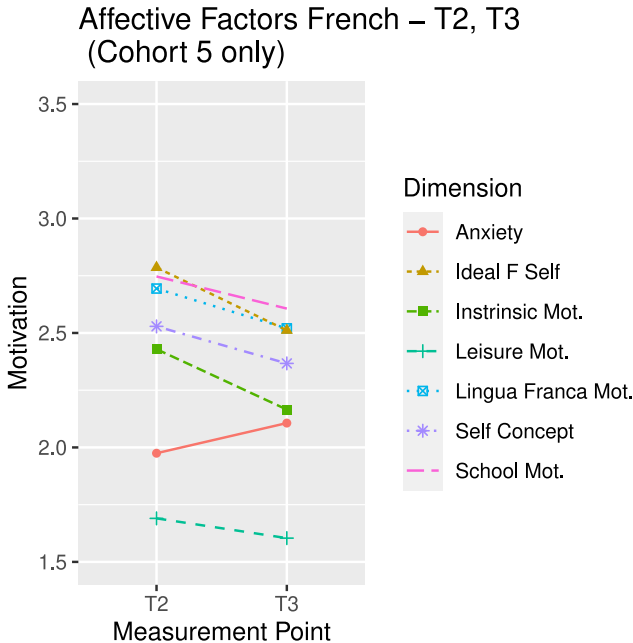


Figure 2: Dynamics of affective dispositions related to French

steadily strengthens for cohort 5, it grows slightly for cohort 4 between the first and second measurement, before dropping again by T3.

While a certain between-subject variation could be identified, the analysis revealed no class effects (therefore, no standard deviations are displayed in the respective column of Table 3).

With regard to French as the second foreign language, Table 3 and Figure 2 suggest a rather different situation. L2 motivation and its antecedents seem less stable than in English. In addition, students generally start with lower motivation and self-concepts at T2, and, at the same time, foreign language anxiety is somewhat higher than in English. Over the course of one school year, students' motivation and self-concepts weaken, whereas anxiety grows notably.

5.2 Teachers and parents

Analogous to the motivation scales, an identical questionnaire for teacher and parental encouragement was administered in both English (T1–T3) and French (T2 and T3, cohort 5 only). Table 4 summarises all effects in relation to perceived teacher and parental encouragement in the two target languages. Fixed effects are plotted in Figures 3 and 4. No interactions between group and measurement point were identified.

Table 4: Fixed and random effects for perceived teacher and parental encouragement. EN: English, FR: French, TE: Teacher encouragement, PE: Parental encouragement, Int: Intercept.

	Int.	Fixed effects (\pm SE)			Random effects	
		Measurement point		Group cohort 5	Student ^a	Class ^b
		T2	T3		SD	SD
EN TE	3.27	-0.08 ± 0.03	-0.30 ± 0.03	0.06 ± 0.08	0.34	0.25
EN PE	2.95	-0.12 ± 0.04	-0.27 ± 0.04	0.02 ± 0.05	0.50	
FR TE	2.97		-0.47 ± 0.05		0.51	0.30
FR PE	2.71		-0.13 ± 0.05		0.59	

^a*n* = 578

^b*n* = 32

Our analysis revealed similar trajectories for both cohorts and target languages with regard to teacher and parental encouragement. While students in

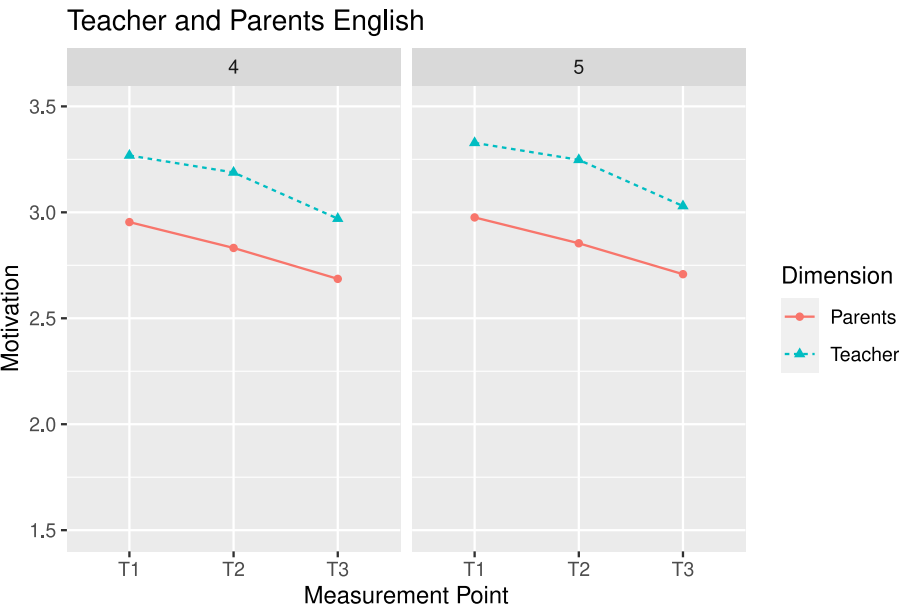


Figure 3: Dynamics of perceived teacher and parental encouragement (English)

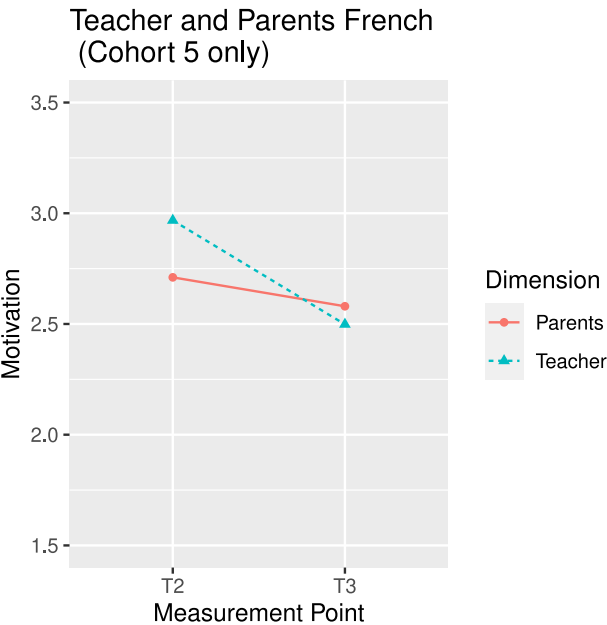


Figure 4: Dynamics of perceived teacher and parental encouragement (French)

both cohorts perceive their L2 teacher and parents as very supportive and encouraging at the first measurement point, this perception weakens over time. In particular, this applies to the French teachers, where the strongest drop was observed. Additionally, considerable between-class variation in both languages was revealed in relation to perceived encouragement by the L2 teacher (cf. analysis report, p. 26: <https://osf.io/r5ypz/>).

5.3 Gender effects

On the basis of previous studies that suggest differences between girls and boys related to the development of foreign language anxiety and the Ideal L2 self, these two dimensions were subjected to further analyses and tested for potential gender effects.

For both target languages, no gender effects were observed with regard to the Ideal L2 self. In contrast, models related to anxiety with gender as an additional fixed effect revealed that girls have higher levels of anxiety than boys. The effect was slightly stronger for English ($\beta_{\text{English}} = 0.18$, $SE = 0.06$; $\beta_{\text{French}} = 0.17$, $SE = 0.09$). At the same time, gender did not interact with measurement point, which indicates that girls' anxiety remains constantly higher than boys'. As pointed out by an anonymous reviewer, there could also be a reporting bias at play, with girls being more likely to admit to feeling anxious than boys.

6 Discussion

With reference to the research question stated in Section 4 above, our results globally support previous findings in that primary school students' motivation and related affective dispositions do not change drastically from middle to late primary school. Accordingly, the most obvious difference was not observed between different measurement points, but between the target languages: English motivation and self-concepts were both higher and more stable, whereas anxiety was lower than in French. Particularly, this applies to the intrinsic motivation, where our results are not entirely in line with previous findings (cf. Heinzmann 2013, Brühwiler & Le Pape Racine 2017). Whereas intrinsic reasons for foreign language learning remain rather positive for English, they are comparatively low for French and even weaken by the end of primary school.

Another result worth highlighting concerns L2 anxiety: Whereas it does not change for English, it intensifies for French. This contradicts previous findings

which observed a decline of perceived stress in French classes over time (cf. Singh & Elmiger 2017). With regard to gender effects, our data support findings by Dewaele et al. (2016) and indicate constantly higher levels of anxiety for girls than for boys in both French and English classes.

Low motivation levels were also identified with regard to the use of English and French in the students' leisure time. The results suggest that pupils do not see much use of these languages for computer games, consulting information on the internet, or understanding the lyrics of their favourite music. The reason therefore could be related to the fact that English and French computer games, songs, or websites are not highly important for young learners at the beginning of their foreign language learning process. This is in line with findings from Heinzmann (2013).

A similar development pattern for both languages was also observed for school-related extrinsic motivation and perceived teacher and parental encouragement, which were rather high at the beginning, but decreased somewhat over time in French as well as in English. This might be due to the fact that pupils become more and more independent of adults and learning languages in order to profit academically becomes less important. In addition, the only notable between-class variation was detected for perceived teacher support. This underpins the critical role of the teacher which is also discussed in Chapters 3 and 7.

Encouraging results could be identified with respect to the ideal L2 self and the lingua franca motivation. The values are particularly high in English, but they are positive and stay above the scale mean in French, too. This supports findings by Heinzmann (2013) and suggests that children believe that they will become competent users of these languages and that the direct application of their language skills in order to communicate with English- or French-speaking people is strongly and steadily endorsed. Contrary to Henry (2009), our analyses revealed no gender effect related to ideal L2 selves. However, Henry's study gives reason to expect that such effects might become apparent with increasing age, as the ideal L2 self is not yet stable at an early age and is expected to become more pronounced over time (cf. also Dörnyei 2009).

7 Conclusion

Our study is based on a quantitative analysis of longitudinal questionnaire data from a relatively large sample. It would be insightful to complement our findings with qualitative measures (such as semi-structured interviews and observations)

in a mixed-methods approach. Furthermore, collecting data at critical moments in the learning process (such as shortly before starting, and at transitions between school types) would provide an additional perspective. Considering these aspects in future studies would allow for a more thorough understanding of how pupils' affective dispositions develop.

These restrictions considered, some general conclusions can be drawn from the results presented in this chapter. Firstly, while there are no drastic changes, primary school pupils have much more promising motivation profiles with regard to English than with regard to French: English learning motivation is stronger, self-concepts are higher, and anxiety is lower than in French. Additionally, all these affective dispositions are more stable in English than in French, where they drop somewhat between the end of the first and second year of instruction. In contrast, the development of motivation regarding academic success, and perceived teacher and parental support proved to be independent of the target language. Our data suggest similar declines in these dimensions for both English and French. An interesting pathway for future studies related to contextual effects on young learners' L2 motivation would be to focus on the perception of peer influence and assess to what extent individual students in a classroom affect their peers' L2 motivation and self-concepts.

References

- Bader, Ursula & Ursula Schaer. 2005. *Evaluation Englisch in den 6. Klassen Appenzell Innerrhoden 2005*. Aarau: PH FHNW.
- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: 10.18637/jss.v067.i01.
- Brühwiler, Christian & Christine Le Pape Racine. 2017. Entwicklung motivationaler Orientierungen beim multiplen Erwerb von Französisch und Englisch als Fremdsprachen am Übergang von der Primarstufe in die Sekundarstufe I. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 22(1). 167–181.
- Busse, Vera. 2017. Plurilingualism in Europe: Exploring attitudes toward English and other European languages among adolescents in Bulgaria, Germany, the Netherlands, and Spain. *The Modern Language Journal* 101(3). 566–582. DOI: 10.1111/modl.12415.
- Buyl, Aafke & Alex Housen. 2014. Factors, processes and outcomes of early immersion education in the francophone community in Belgium. *International Journal of Bilingual Education and Bilingualism* 17(2). 178–196. DOI: 10.1080/13670050.2013.866627.

- Cenoz, Jasone. 2004. Teaching english as a third language: the effect of attitudes and motivation. In Charlotte Hoffmann & Jehannes Ytsma (eds.), *Trilingualism in family, school, and community*. Clevedon: Multilingual Matters.
- Dewaele, Jean-Marc, Peter Macintyre, Carmen Boudreau & Livia Dewaele. 2016. Do girls have all the fun? anxiety and enjoyment in the foreign language classroom. *Theory and Practice of Second Language Acquisition* 2. 41–63.
- Dörnyei, Zoltán. 2009. The L2 motivational self system. In Zoltán Dörnyei & Ema Ushioda (eds.), *Motivation, language identity and the L2 self*. Bristol: Multilingual Matters.
- Dörnyei, Zoltán & István Ottó. 1998. Motivation in action: A process model of L2 motivation. *Working Papers in Applied Linguistics* 4. 43–69.
- Dörnyei, Zoltán & Ema Ushioda. 2011. *Teaching and researching motivation*. 2nd edn. Harlow: Longman.
- Gardner, Robert C. 2006. The socio-educational model of second language acquisition: a research paradigm. *EUROSLA Yearbook* 6. 237–260.
- Gardner, Robert C. 2007. Motivation and second language acquisition. *Porta Linguarum* 8. 9–20.
- Graham, Suzanne, Louise Courtney, Alan Tonkyn & Theodoros Marinis. 2016. Motivational trajectories for early language learning across the primary–secondary school transition. *British Educational Research Journal* 42(4). 682–702. DOI: 10.1002/berj.3230.
- Heinzmann, Sybille. 2010. Hat die einföhrung von "fröhenglisch" in der primarschule einen einfluss auf die motivation der primarschulkinder französisch zu lernen? *Bulletin VALS-ASLA* 91. 7–27.
- Heinzmann, Sybille. 2013. *Young language learners' motivation and attitudes: Longitudinal, comparative and explanatory perspectives*. London: Bloomsbury.
- Henry, Alastair. 2009. Gender differences in compulsory school pupils' L2 self-concepts: A longitudinal study. *System* 37(2). 177–193.
- Hufeisen, Britta. 2003. Kurze Einföhrung in die linguistische Basis. In Britta Hufeisen & Gerhard Neuner (eds.), *Mehrsprachigkeitskonzept – Tertiärsprachen – Deutsch nach Englisch*, 7–11. Strasbourg: Council of Europe Publishing.
- MacIntyre, Peter D. 1999. Language anxiety: A review of the research for language teachers. In Dolly Jesusita Young (ed.), *Affect in foreign language and second language teaching: A practical guide to creating a low-anxiety classroom atmosphere*, 24–45. Boston, MA: McGraw-Hill.
- McGroarty, Mary. 2001. Situating second language motivation. In Zoltán Dörnyei & Richard Schmidt (eds.), *Motivation and second language acquisition*, 69–90. Honolulu: University of Hawaii Second Language Teaching & Curriculum Center.

- Meissner, Franz-Joseph, Christine Beckmann & Anna Schröder-Sura. 2008. *Mehrsprachigkeit fördern. vielfalt und reichum europas in der schule nutzen* (MES). Giessen: Universität Giessen.
- Mihaljević Djigunović, Jelena. 2012. Attitudes and motivation in early foreign language learning. *Center for Educational Policy Studies Journal* 2. 55–74.
- Mihaljević Djigunović, Jelena & Lucilla Lopriore. 2011. The learner: Do individual differences matter? In Janet Enever (ed.), *ELLiE: Early language learning in Europe*, 43–60. London: The British Council.
- Mihaljević Djigunović, Jelena & Marianne Nikolov. 2019. Motivation of young learners of foreign languages. In Martin Lamb, Kata Csizér, Alastair Henry & Stephen Ryan (eds.), *The Palgrave handbook of motivation for language learning*, 515–533. Cham: Springer. DOI: 10.1007/978-3-030-28380-3_25.
- Nikolov, Marianne. 2002. *Issues in english language education*. Bern: Peter Lang.
- Pfenninger, Simone & David Singleton. 2016. Affect trumps age: a person-in-context relational view of age and motivation in SLA. *Second Language Research* 32(3). 311–345. DOI: 10.1177/0267658315624476.
- R Core Team. 2019. *R: A language and environment for statistical computing*. Version 3.5.1. Wien, Österreich. <https://www.R-project.org/>.
- Schaer, Ursula & Ursula Bader. 2003. *Evaluation Englisch an der Primarschule. Projekt 012. Appenzell Innerrhoden. Mai bis November 2003*. Zürich, Luzern: PHZH, PHZ Luzern, ILLB Bern.
- Singh, Lisa & Daniel Elmiger. 2017. *Externe Evaluierung der Pilotphase des Französisch- und Englischunterrichts nach Passepartout: Schuljahre 2009–2016*. Neuchâtel: IRDP Institut de recherche et de documentation pédagogique.
- Stöckli, Georg. 2004. *Motivation im Fremdsprachenunterricht: Eine theoriegeleitete empirische Untersuchung in 5. und 6. Primarschulklassen mit Unterricht in Englisch und Französisch* (Pädagogik bei Sauerländer). Oberentfelden: Sauerländer Verlag.
- Tragant, Elsa. 2006. Language learning motivation and age. In Carmen Muñoz (ed.), *Age and the rate of foreign language learning*, 237–276. Clevedon: Multilingual Matters.
- Ushioda, Ema. 2017. The impact of global english on motivation to learn other languages: toward an ideal multilingual self. *The Modern Language Journal* 101(3). 469–482. DOI: 10.1111/modl.12413.
- Winter, Bodo. 2013. *Linear models and linear mixed effects models in R with linguistic applications*. <http://arxiv.org/pdf/1308.5499.pdf> (20 September, 2018).

Chapter 9

Language aptitude in German as a school language and English as a foreign language in primary school

Hansjakob Schneider


Zurich University of Teacher Education

Two cases of language learning (English as a foreign language and German as the language of instruction) were compared using data from the LAPS II project. Hierarchical Regression Analyses were performed with achievement in English or German respectively as dependent variables and language aptitude components as well as other cognitive and language-related variables as independent variables. Results suggest that language aptitude variables are strong predictors of achievement in English as a foreign language and to a lesser degree for German as the language of schooling. Language self-concepts play an important role in achievement in English and less so in German. In the case of German (but not of English), socioeconomic status is a strong predictor of achievement. Since these analyses were exploratory, further research is needed to come to conclusions relevant for language teaching.

1 Introduction

In the previous chapters the focus has been on English or French as foreign languages. Language aptitude is indeed a concept which in research is more often linked with foreign language learning rather than untutored second language acquisition or first language acquisition (untutored in early childhood or tutored in school). The present chapter explores the relationship between language aptitude and achievement in German as a school language (GSCL) and compares it to the case of English as a Foreign Language (EFL).



Hansjakob Schneider. 2021. Language aptitude in German as a school language and English as a foreign language in primary school. In Raphael Berthele & Isabelle Udry (eds.), *Individual differences in early instructed language learning: The role of language aptitude, cognition, and motivation*, 179–195. Berlin: Language Science Press. DOI: [10.5281/zenodo.5464761](https://doi.org/10.5281/zenodo.5464761) 

The objective of this chapter is to find out if achievement in EFL and in GSCL share common influencing factors and to identify them. The role of aptitude in language learning and its position in the field of potentially influential other factors is of particular interest. In this respect the quantitative-analytical procedure employed is exploratory.

In §2 theoretical and empirical findings about the role of language learning aptitude are discussed. The study design is presented in §3. §4 contains statistical analyses: Results on the achievement in EFL and GSCL are presented and variables with a high influence on the achievement in these languages are identified. The results are discussed in §5 in view of the theory of language aptitude.

The analyses used are similar to the ones in Chapter 4 and yet different in specific ways. The main focus of the research presented in this volume is on foreign language learning and the data we collected pertain mainly to this matter. This includes ID variables beyond aptitude for EFL (e.g. different kinds of motivation). In Chapters 3 and 4, EFL was conceptualized as the dependent variable, while achievement in GSCL was designed to be one of the independent variables. Since EFL had a special status in the LAPS II study, some language-specific data were collected for EFL only and not for German. Therefore, the choice of variables to be compared was restricted to variables not specific to either language (working memory for instance) and to self-concept (which was collected for English as well as for German). For a full account of the situation for EFL see Chapters 3 and 4.

2 Aptitude and related variables in EFL- and GSCL-learning

For foreign language learning in school there is plenty of evidence that aptitude has significant impact on the learning outcome (Li 2015). Studies which investigate language aptitude in the acquisition of first languages and especially in learning first languages as the languages of schooling on the other hand are rather scarce (but see e.g. Skehan & Ducroquet 1988). Generally the impact of aptitude on learning the language of instruction in school is less easily proven, maybe because the “massive exposure time and experience with native languages overrides genetic influences and “levels them out” – influences and differences which would potentially have been there in the first place as well.” (Reiterer 2018: ix)

Learning the language used in the general school-context (language of instruction) is an interesting case of language acquisition and learning: On the one hand,

it relies heavily on first language acquisition (at least in situations where it is identical or related to the first language of the areas in question). On the other hand, in school the basic skills of reading and writing are introduced and fostered; moreover, with the *language of schooling* (Schleppegrell 2004) features of language which can be quite different from everyday language use come into play. The few existing studies about untutored vs. tutored L2 learning/acquiring present contradictory results as far as the role of language learning aptitude is concerned (Udry et al. 2019).¹ In Li's meta-analysis of research on associations between aptitude and grammar learning, aptitude showed significant effect sizes also in non-tutored contexts; however, Li views these results critically because the studies reviewed lack control over the acquisition contexts of the participants (Li 2015: 405). There is, however, empirical support for the thesis that language learning aptitude plays a role not only in foreign language learning/acquisition but also in first language acquisition (Biedroń & Pawlak 2016).

Beyond language learning aptitude there is a number of variables which have proved to contribute significantly to L2- and L1-learning (for a detailed account see Chapter 1). For the purposes of this chapter I will present only the ones for which data were collected for both EFL and GSCL (for detailed information see Chapter 1).

Nonverbal fluid intelligence is related closely to learning in general, to language aptitude and hence to L2 learning as well (see Chapter 1); it does correlate with aptitude, but can be differentiated from it (Sparks et al. 2012). Fluid intelligence is also a moderate predictor of school-related abilities in L1 (such as reading comprehension, Peng et al. 2019).

Another nonlinguistic variable is working memory. Working memory, like nonverbal intelligence, is a concept with impact on many forms of learning. Its influence on L2 learning in the context of aptitude studies has been empirically proven (Wen et al. 2017, see also Chapter 1). Working memory also influences L1 reading comprehension at school (e.g. Peng et al. 2017).

Competence in L1 has been recognized in recent years as an important variable for foreign language learning. Sparks and his colleagues have put forward their linguistic coding difference hypothesis (LCDH) and have gathered sound empirical evidence to show that attainment in L1 predicts learning outcomes in L2 on secondary school level (Sparks et al. 2012, Li 2016).

In a similar vein, in a longitudinal study (grades 1–10) Sparks et al. (2012) determined the effect sizes of L2 aptitude, L1 skills, IQ and other variables on various facets of L2 proficiency using Hierarchical Regression Analyses. They found

¹No studies comparing the influence of language aptitude in the cases of untutored acquisition vs. learning of L1 at school are known to me.

strong influences of a composite of L1 skills, IQ and aptitude on all areas of L2 proficiency; in the case of reading comprehension in L2 the variance accounted for by this composite was 45%. Unfortunately, the unique roles of aptitude and IQ were not reported separately.

The role of language self-concept in L2-learning is reported in Chapters 1, 3, 4 and 8. As for reading comprehension in GSCL there is evidence from a large longitudinal study for the skill-development hypothesis (achievement predicts self-concept) as well as for reciprocal effects (in addition to skill-development hypothesis: self-concept predicts achievement, especially around grade 5, Retelsdorf et al. 2014).

The impact of socioeconomic status (SES) on school achievement in general is well documented (see the meta-analysis by Sirin 2005). For its influence on EFL in the LAPS project see Chapter 5. Its status for reading comprehension in GSCL is undisputed: A vast number of empirical studies have proven a strong influence of SES on reading comprehension in school contexts for all grades (Schaffner 2009).

The conclusion of all these findings is that both for EFL and GSCL learning is a multivariate process. But is the influence of aptitude (in the context of the above-mentioned variables) on learning EFL similar to learning GSCL? In other words: Do aptitude variables influence tutored L1- and L2-learning in a similar way or is there a difference in strength (effect size) and kind of aptitude components (e.g. grammatical sensitivity vs. inductive ability)? From these questions the following objective can be formulated:

The main objective of the current article is to quantify the effect of language aptitude on GSCL in comparison to EFL.

3 Study design and method

The data to be analysed here stem from the LAPS II study. Since the general design of the study is described in Chapter 2, I will only go into aspects of the design which are specific to the research objective mentioned at the end of the last section.

The data come from the LAPS II sample ($n(T2) = 578$, $n(T3) = 566$, cf. §4.5). The sample size in this chapter is not identical because only students were included for whom data for the two measurement points and for all variables in question were available. For this reason, the number of participants varies depending on the type of analysis carried out. Furthermore, outliers and students with English as L1 were excluded.

In the following, the two age groups, 1 (grade 4 to 5) and 2 (grade 5 to 6), are treated separately for reasons to be presented in §4. Whereas Chapters 3 and 4 in

this volume are concerned with understanding the underlying dimensions of EFL (Chapter 3) and predicting achievement in EFL (Chapters 4 and 5), a different approach is taken in the present chapter: The main idea of the following analyses is to compare the influence of language aptitude on performance in EFL and GSCL respectively, taking into account variables collected for both languages. On the one hand, these are variables of nonlinguistic nature (e.g. general cognitive variables, socioeconomic status), on the other hand, (cognitive-linguistic) aptitude variables. Furthermore, the respective self-concepts for EFL and GSCL are also included in the analyses.

Excluded are all variables which are not available for both languages. These are mainly the variables of language learning motivation (only available for English). It seems obvious that a comparison of influencing variables in EFL and GSCL should draw on the same or similar variables. As mentioned in the introduction, the consequence of this procedure is that the results cannot be compared to similar analyses of EFL alone in Chapters 3 and 4.

Hierarchical regression models were fitted to the data to identify the variables that predict performance in English and German. The following variables were included in the analyses, achievement in English and German being the dependent variables (For a more detailed description see Chapter 2):

Achievement in English as a foreign language (EFL): This variable was measured using the Oxford Young Learners Proficiency Test (OYLPT) for T1 and a c-test for T2 and T3. Since the OYLPT does not measure the same linguistic dimensions as the c-tests (it includes a listening comprehension part, for instance) only measurement point T3 will be included in the analysis of predicting variables (cf. §4.2 for further details). C-tests have proved to be a valid measure for general language proficiency (Eckes & Grotjahn 2006).

Achievement in German as the school language (GSCL): For all three measurement points the sentence reading part of the ELFE test (Lenhard & Schneider 2006) was administered. ELFE is a widely used standardised test covering grades 1 to 6. Within narrow time limits students are confronted with a series of incomplete sentences. In each sentence they must choose one of 5 words to form a correct sentence; a translated example would be:

Lea is playing (after/next to/**instead**/then/that) of learning.

The test addresses primarily the dimensions of vocabulary and of grammar. In this sense it covers similar dimensions as the c-tests used for measuring achievement in EFL (but not the productive component of coming up with a suitable word). The measure was the number of sentences completed correctly per minute.

Grammatical sensitivity (MLAT, Part 2): This test is a translated and adapted version of the MLAT-E, Part 2 (Carroll & Sapon 2010).

Inductive ability (PLAB): For inductive ability a translated and adapted version of part 4 of the PLAB (Pimsleur et al. 2004) was used.

Phonemic coding ability (Llama): We used Llama-D, the phonemic discrimination task in which students had to recognize sounds (Meara et al. 2001).

Verbal working memory (backwards digit span): Verbal working memory was tested with a task in which students had to memorize a sequence of numbers (increasing in length) in backward order.

Visuospatial working memory (Corsi block task): Visuospatial working memory was tested using an adaptation of the Corsi block task. Students had to remember the order of an increasing number of squares lit up from a matrix of squares.

Field independence (group embedded figures test): Participants had to find simple geometrical figures embedded in more complex figures under time pressure (Oltman et al. 1971).

Fluid intelligence (CFT): Students had to discover rules for example in comparing sequences of increasingly complex matrices (Weiß 2006).

Socioeconomic status, economic and cultural aspects (SES): In line with the theoretical assumptions and empirical findings in Chapter 5, the construct of the SES is split up into two dimensions: economic and cultural aspects (for the variables involved in each dimension see Chapter 5).

The data were statistically analysed using SPSS (version 27 for Macintosh). Procedures included repeated measurement ANOVAS (used to determine whether the two age groups behave similarly) as well as hierarchical stepwise regression analyses (used to establish the kind and strength of influencing variables).

4 Results

4.1 Achievement in German as a School Language and English as a foreign language

Since the sample of the longitudinal LAPS II study consisted of two age groups (grades 4 to 5; grades 5 to 6) it had to be established first if the development in

achievement of English (and of German) was homogeneous for the two groups. As shown in Chapter 4 the variable age group (grade 4 or 5 respectively at T1) contributes to attainment in English at T3. There seems to be good reason to assume that the two age groups behave differently as far as learning EFL and other variables mentioned are concerned. To get a more detailed view of the role the two age groups play, a repeated measures variance analysis with the three measurements of English (Figure 1) and of German (Figure 2) as dependent and the age group (1 and 2) as the independent variables was performed. Figures 1 and 2 show that the (younger) age group 1 develops more positively than age group 2.

Z-standardized values were used to make the effects more visible. In Figure 1 age group 1 not surprisingly has relatively lower mean measures for achievement in EFL than age group 2 (students of age group 2 having profited of one more year of English lessons). Z-transformed values can be negative (as is the case for all means of age group 1). This means that they are below average of the whole sample (the average of the whole sample being standardized at the value 0).

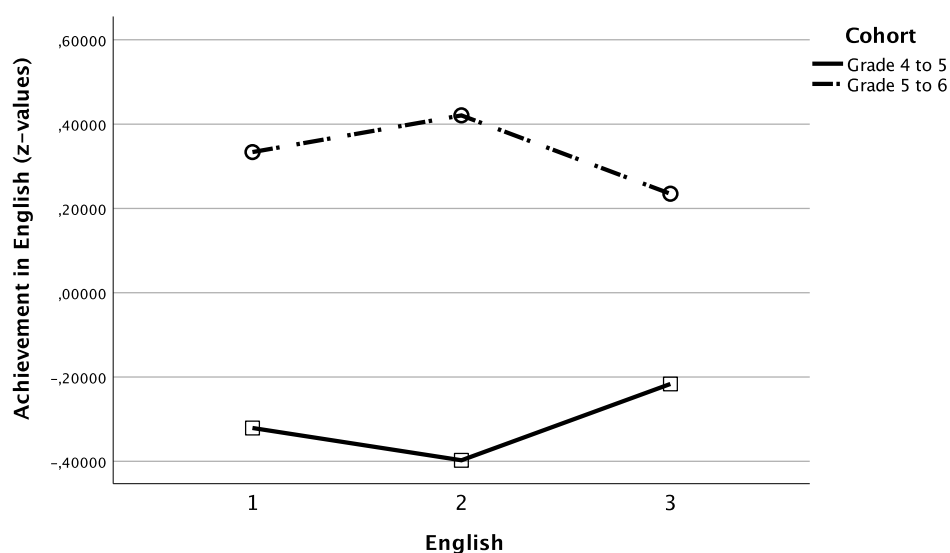


Figure 1: The development of test results in English (standardized) from T1 to T3 for the two age groups (grades 4 to 5, grades 5 to 6); Repeated Measures ANOVA: within subjects (measurements T1–T3)*between subject (age groups) effects significant (Greenhouse-Geisser, $F = 14.37$, $df = 1.65$, $p < 0.001$); covariates: fluid intelligence CFT 20 ($F = 45$, $p < 0.001$); SES ($F = 13.44$, $df = 1$, $p < 0.001$)

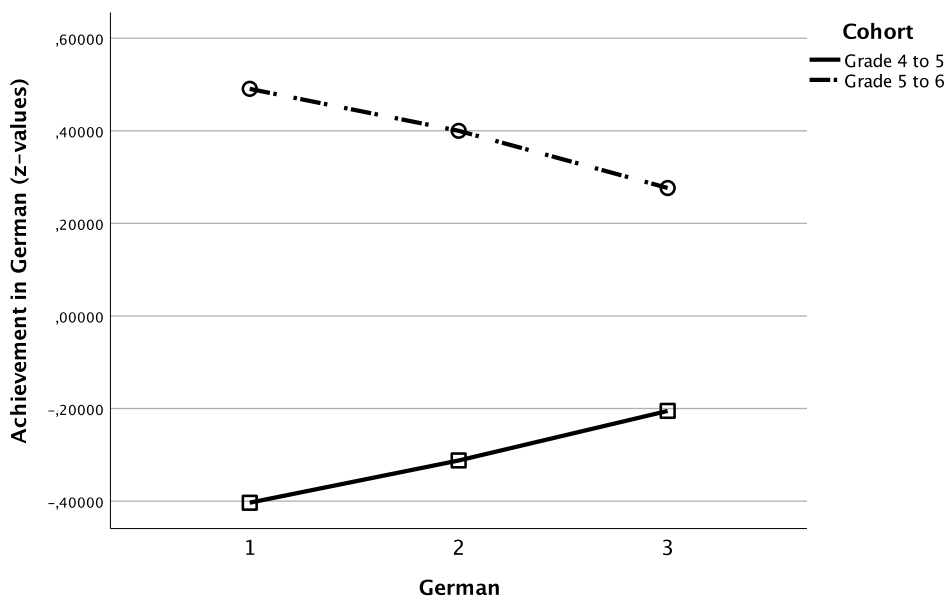


Figure 2: The development of test results in German (standardized measures of the ELFE-subtest on sentence reading) T1 to T3 for the two age groups (grades 4 to 5, grades 5 to 6); Repeated Measures ANOVA: within subjects (measurements T1–T3)*between subject (age groups) effects significant (Greenhouse-Geisser, $F = 22.53$, $df = 1.94$, $p < 0.001$); covariates: fluid intelligence CFT 20 ($F = 64.86$, $p < 0.001$); SES ($F = 52.76$, $df = 1$, $p < 0.001$)

In both EFL and GSCL the difference in achievement between the two age groups diminishes with time. This means that towards the end of primary school (between grades 5 and 6) learning in EFL as well as in GSCL slows down whereas between grade 4 and 5 (age group 1, T1–T3) more learning takes place. The repeated measures ANOVAs reported in Fig. 1 and 2 prove that the interaction between the age groups and the development of achievement in either language is significant, i.e. that the two groups develop differently in English and in German. These results led to the decision to perform further analyses separately for the two age groups.

Another element which links the development of EFL and GSCL is stability over time: although in absolute terms the students' test results increase over the period of almost two years, they are highly correlated for the three measurement points. Tables 1 and 2 show the Pearson correlations of the test results for the English and the German test respectively.

Table 1: Pearson correlations of test results in English for age groups 1 and 2 from T1 to T3; ***: $p < 0.001$, age group 1 $n = 252$, age group 2 $n = 286$.

		Age group 1			Age group 2		
		E T1	E T2	E T3	E T1	E T2	E T3
E T1	1		0.67***	0.65***	1	0.72***	0.71***
E T2			1	0.84***		1	0.88***
E T3				1			1

Table 2: Pearson correlations of test results in German for age groups 1 and 2 from T1 to T3; ***: $p < 0.001$, age group 1 $n = 260$, age group 2 $n = 283$.

		Age group 1			Age group 2		
		G T1	G T2	G T3	G T1	G T2	G T3
G T1	1		0.79***	0.74***	1	0.80***	0.77***
G T2			1	0.76***		1	0.75***
G T3				1			1

The test results are so strongly correlated that the assumption of common underlying factors (one for the three English measures and one for the three German measures) lends itself. Similarly, both aptitude variables used in our study (grammatical sensitivity and inductive ability) are shown in Chapter 10 to be remarkably stable over time.

In the next section the impact of language aptitude on achievement in English and in German is analysed.

4.2 Influencing variables on German as a school language and English as a foreign language

One of the interesting questions in the theory of language aptitude pertains to the specificity of the aptitude concept for foreign language learning, second language acquisition, first language acquisition and school language learning. Of these four different situations, the LAPS II project is able to compare foreign language learning with school language learning (which is partly connected to first

language acquisition in that it builds upon the latter but introduces novel dimensions like reading or writing, see §2). The variables included in the statistical model can be grouped into three sections:

1. Language-learning specific
 - a. Grammatical sensitivity (MLAT)
 - b. Inductive ability (PLAB)
 - c. Phonemic encoding (Llama)
 - d. Self-concept for English/German
2. General cognitive
 - a. Working memory (Backward digit span, Corsi blocks)
 - b. Field independence (GEFT)
 - c. Nonverbal fluid intelligence (CFT)
3. Sociodemographic
 - a. Socio-economic status (economic)
 - b. Socio-economic status (cultural)

The following analyses use multiple stepwise regression analyses to establish the nature and effect size of influencing variables on the achievement in English and in German. The aim is to compare the patterns found for achievement in English and in German. In contrast to Chapter 4 the achievement measures of the two languages at T1 are not included as predictors of T3 for the following reason: The English test at T1 (OYLPT) is not the same as at T3 (c-test), whereas for German it is (ELFE). A comparison of the patterns of influencing variables for the two languages would be affected by this inequality.

In comparing the predicting variables for achievement in English and in German the respective measures for T3 (i.e. English c-test T3 and ELFE test T3) were chosen as the dependent variables. Of the language-learning-specific variables mentioned above, the measures for T1 were entered into the analyses as predictors for the T3 outcome variables. Since the nature of these analyses is exploratory, the stepwise method is applied, which means that the independent variables selected in the model are chosen not for theoretical reasons but purely for their ability to contribute significantly to explaining variance (Field 2009: 213). The strongest variable is chosen first, the second variable is the one which explains most of the remaining variance and so forth. All variables mentioned

above were included in the models but only the variables leading to significant changes in explained variance (R^2) are reported in the tables. Tables 3 and 4 show the results of these analyses for EFL (age groups 1 and 2 respectively).

Table 3: Grade 5 – results of the stepwise multiple regression analysis for achievement in English (c-test T3) with predictors T1, ***: $p < 0.001$, **: $p < 0.01$, *: $p < 0.05$; $n = 170$. SCE: Self-concept English, IA: Inductive ability, FI: Fluid intelligence, BDS: Backward digit span, PCA: Phonemic coding ability

	R^2	Corrected R^2	Change in R^2	Change in F
<i>Model 1: SCE T1</i>	0.23	0.23	0.23	51.00***
<i>Model 2: Model 1 + IA T1</i>	0.33	0.32	0.10	23.91***
<i>Model 3: Model 2 + FI T1</i>	0.37	0.36	0.04	10.91**
<i>Model 4: Model 3 + BDS T1</i>	0.40	0.38	0.03	6.68*
<i>Model 5: Model 4 + PCA T1</i>	0.41	0.39	0.02	4.10*

Table 4: Grade 6 – results of the stepwise multiple regression analysis for achievement in English (c-test T3) with predictors T1; *** $p < 0.001$; $n = 223$. GS: Grammatical sensitivity, SCE: Self-concept English, IA: Inductive ability.

	R^2	Corrected R^2	Change in R^2	Change in F
<i>Model 1: GS T1</i>	0.26	0.26	0.26	78.59***
<i>Model 2: Model 1 + SCE T1</i>	0.39	0.38	0.12	44.05***
<i>Model 3: Model 2 + IA T1</i>	0.42	0.42	0.04	14.26***

The models for the two age groups are similar in the sense that their strongest predictor variables are comparable: self-concept English and components of aptitude. In the final model of grade 5, general fluid intelligence and verbal working memory have minor influence as well. Interestingly, fluid intelligence, a strong predictor of achievement at school in general, seems to lose some of its predictive power to the more language specific language aptitude variables. The variables that turn out to be most predictive in these models belong to the ones that have been shown to be positively associated with foreign language achievement in Chapters 3 and 4.

The models of the two age groups for achievement in German (cf. Tables 5 and 6) are also similar in that the respective final models include the same predictors: language aptitude (grammatical sensitivity), self-concept German, and

SES (cultural aspects). Fluid intelligence, the strongest predictor for younger students (grade 5) does not show a significant effect in the older age group (grade 6). For the older students grammatical sensitivity is the leading predictor of achievement in German.

Table 5: Grade 5 – results of the stepwise multiple regression model for achievement in German (ELFE T3) and predictors T1; *** $p < 0.001$, ** $p < 0.01$, $n = 170$. FI: Fluid intelligence, SESc: SES (cultural), SCG: Self-concept German, GS: Grammatical sensitivity.

	R^2	Corrected R^2	Change in R^2	Change in F
<i>Model 1: FI T1</i>	0.16	0.16	0.16	32.54***
<i>Model 2: Model 1 + SESc T1</i>	0.23	0.23	0.07	15.74***
<i>Model 3: Model 2 + SCG T1</i>	0.27	0.26	0.03	7.79**
<i>Model 4: Model 3 + GS T1</i>	0.29	0.27	0.02	4.60*

Table 6: Grade 6 – results of the stepwise multiple regression model for achievement in German (ELFE T3) and predictors T1; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, $n = 222$

	R^2	Corrected R^2	Change in R^2	Change in F
<i>Model 1: GS T1</i>	0.17	0.16	0.17	43.33***
<i>Model 2: Model 1 + SESc T1</i>	0.21	0.20	0.04	11.24**
<i>Model 3: Model 2 + SCG T1</i>	0.23	0.22	0.02	5.63*

In conclusion we can state that for English and also for German (and for both age groups) aptitude variables (inductive ability, grammatical sensitivity, and phonemic coding ability in English; grammatical sensitivity in German) play a key role in accounting for achievement. They explain more unique variance for EFL (12% and 30%) than for GSCL (between 2% and 17%). SES seems to be more important in learning GSCL than EFL. Moreover, general fluid (nonverbal) intelligence accounts for variance in the models for the younger students but not for the older ones. It seems that nonverbal intelligence is replaced as a predictor by the more language specific aptitude measures. Finally, language self-concept is a relatively stronger predictor for achievement in EFL than in GSCL.

5 Discussion

The main objective of the present chapter was to quantify the effect of language aptitude on EFL and GSCL. The short conclusion is this: Of all the language-learning specific, general cognitive and sociodemographic measures, language learning aptitude (mainly in the facets of grammatical sensitivity and inductive ability) and self-concept overrule the other variables in the case of achievement in English. For German, the influence of language aptitude variables is weaker and it differs between the two age groups. In the case of EFL, the strongest aptitude variable for the younger students is inductive ability, for the older students grammatical sensitivity takes the lead. In the case of GSCL, grammatical sensitivity is the strongest aptitude variable for both age groups.

The more detailed conclusion focuses on two aspects: (a) The stability over time of some of the constructs presented, (b) structural differences between learning EFL vs. GSCL.

(a) The high correlations between the achievement measures in English and German over the period from grades 4 to 6 suggest that while achievement increases during this period (i.e. the students improve intra-individually), there seems to be one underlying ability for EFL and one for GSCL (there is not much inter-individual variation in either language, high and low achievers for example remain high and low over time). In the same sense, but to a slightly lesser degree, grammatical sensitivity and inductive ability are stable over time (cf. Chapter 10).

(b) The structure and strength of the influencing variables were shown to be very similar for the achievement in English and the achievement in German. The aptitude variables *grammatical sensitivity* and *inductive ability* explain a large part of the variance in the regression models for EFL. This strong influence of language aptitude variables on foreign language learning is well documented (e.g. Li 2016) and not surprising.

The lower influence of language aptitude (in the form of grammatical sensitivity) on the comprehension of German sentences (Tables 5 and 6) may be attributed to differing situations of acquisition: In grades 4 to 6 the predominant part of students has reached an achievement level in German which exceeds their achievement in English by far. As discussed in §2 the influences of language aptitude may have been levelled out by vast experience with the language.

Whereas in EFL inductive ability is a predictor for achievement (especially for the younger age group), in the case of GSCL it seems to play a much weaker part (relative to grammatical sensitivity). This fact could be interpreted along the following lines: In the case of German (as L1) the period of having to induce grammatical rules and structures is long past for most students. Students do not

have to induce grammatical rules because they acquired them in earlier stages (mostly in pre-school acquisition). In contrast, the ability to recognise grammatical functions is an important prerequisite for selecting the correct alternative word in the sentences of the ELFE-test (is the word in the role of an agent, patient, theme, etc.?). Therefore, grammatical sensitivity is a key feature in the context of this test.

In the case of English, the same reasoning holds true: To fill the gaps in the c-tests the function of the words in the sentence must be clear (and in addition morphosyntactic knowledge must be retrieved). Why inductive ability is a stronger predictor than grammatical sensitivity in age group 1 (grade 4 to 5) is difficult to answer. Rule induction might be more important in the early stages of learning a language than in later stages.

Socioeconomic status significantly influences the achievement in German but not in English. This is probably attributable to the abundance of socially embedded experiences with the German language that speakers of German have made up to the age of 10 to 12. In the case of English, direct influence of family background is much shorter (English having only been taught for 2 to 3 years in our sample) and much less intensive (English is not generally spoken in families with German as L1).

Results not reaching statistical significance may be at least as instructive as the statistically significant results discussed so far: One aptitude variable, phonemic coding ability, proved to have only a marginal effect in our study. This can be explained by the nature of the language tests administered. Both for English and for German, written texts (or sentences) were used. It could be argued that in reading as well as in listening phonemic coding plays a role. However, in reading the role of phonemic coding is especially important in the early stages of literacy development when students translate each grapheme into a phoneme and store it in a buffer until the whole word has been synthesized. According to the dual coding theory (cf. Coltheart et al. 2001) another cognitive path becomes more and more dominant as proficiency in reading increases: it leads from orthographic analysis (a graphemic representation of a word) directly to the orthographic lexicon, and from there to the semantic system; the semantic content is then matched to the corresponding phonemic representation as a whole (i.e. without analysing or synthesizing). The role of phonemic coding ability in this route is minimal and therefore the Llama test accounts for hardly any variance in the reading of English or German (though more in early stage EFL than in advanced GSCL). Its position might have been quite different had the tests involved speaking or listening activities. The same goes for verbal working memory: Had the tests involved oral language, verbal working memory might have played a more important role.

The relatedness of intelligence and language aptitude has been discussed in the literature (and empirically proved in Chapter 3) and the upshot is that they are closely related but distinguishable constructs (e.g. Li 2016: 826f.). When using stepwise multiple regression models, this situation may lead to one of two closely competing variables to take the lead (and explain most of the unique variance) leaving its competitor with only a small part of explained variance. This seems to be the case for sentence comprehension in German: For students in grade 5 fluid intelligence is the strongest predictor (Table 5) whereas for the sixth graders grammatical sensitivity takes the lead.

Two final words of caution: Firstly, the tests used for measuring achievement in English (c-tests) and German (sentence comprehension) are similar but not identical (c-tests require productive skills to a higher degree). Some of the differences in the patterns of influencing variables may be attributable to these differences. However, since according to Li's meta-analysis on the influence of language aptitude on language learning the language analytic component is relatively weakly correlated with the productive skills of speaking and writing (Li 2016: 823f.), the difference between the tests may not be all that important for the pattern of grammatical sensitivity and inductive ability in the English and German models.

Secondly, the analyses performed for the purposes of this chapter are exploratory in nature. While they do show the trend for language aptitude variables to be the strongest predictors for achievement in EFL and have some influence in GSCL, it would be too early to draw specific conclusions for teaching at school. Rather, further research should investigate whether aptitude variables can be effectively taught in school settings and what their impact is on various measures of achievement (grammar, listening, reading, speaking, and writing).

References

- Biedroń, Adriana & Mirosław Pawlak. 2016. New conceptualizations of linguistic giftedness. *Language Teaching* 49(2). 151–185. DOI: 10.1017/S0261444815000439.
- Carroll, John B. & Stanley M. Sapon. 2010. *Modern language aptitude test: Elementary manual*. Rockville: Second Language Testing Foundation.
- Coltheart, Max, Kathleen Rastle, Conrad Perry, Robyn Langdon & Johannes Ziegler. 2001. DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review* 108(1). 204–256. DOI: 10.1037/0033-295X.108.1.204.
- Eckes, Thomas & Rüdiger Grotjahn. 2006. A closer look at the construct validity of c-tests. *Language Testing* 23(3). 290–325. DOI: 10.1191/0265532206lt330oa.

- Field, Andy. 2009. *Discovering statistics using SPSS*. London: SAGE Publications Ltd.
- Lenhard, Wolfgang & Wolfgang Schneider. 2006. *ELFE 1-6 Ein Leseverständnistest für Erst- bis Sechstklässler*. Göttingen: Hogrefe.
- Li, Shaofeng. 2015. The associations between language aptitude and second language grammar acquisition: a meta-analytic review of five decades of research. *Applied Linguistics* 36(3). 385–408. DOI: 10.1093/applin/amu054.
- Li, Shaofeng. 2016. The construct validity of language aptitude: A meta-analysis. *Studies in Second Language Acquisition* 38(4). 801–842. DOI: 10.1017/S027226311500042X.
- Meara, Paul M., James Milton & Nuria Lorenzo-Dus. 2001. *Llama language aptitude tests: The manual*. Swansea.
- Oltman, Philip K., Evelyn Raskin, Herman A. Witkin & Consulting Psychologists Press. 1971. *Group embedded figures test*. [Online; accessed 2015-09-05]. Consulting Psychologists Press Palo Alto, CA.
- Peng, Peng, Marcia Barnes, Cuicui Wang, Wei Wang, Shan Li, H. Lee Swanson, William Dardick & Sha Tao. 2017. A meta-analysis on the relation between reading and working memory. *Psychological Bulletin* 144(1). 48–76. DOI: 10.1037/bul0000124.
- Peng, Peng, Tengfei Wang, CuiCui Wang & Xin Li. 2019. A meta-analysis on the relation between fluid intelligence and reading/mathematics: Effects of tasks, age, and social economics status. *Psychological Bulletin* 145(2). 189–236.
- Pimsleur, Paul, Daniel J. Reed & Charles W. Stansfield. 2004. *Pimsleur language aptitude battery PLAB*. Rockville: Second Language Testing.
- Reiterer, Susanne M. 2018. Preface. In Susanne M. Reiterer (ed.), *Exploring language aptitude: views from psychology, the language sciences, and cognitive neuroscience*, vii–xii. Cham: Springer.
- Retelsdorf, Jan, Olaf Köller & Jens Möller. 2014. Reading achievement and reading self-concept – testing the reciprocal effects model. *Learning and Instruction* 29(2). 21–30.
- Schaffner, Ellen. 2009. Determinanten des leseverstehens. In Wolfgang Lenhard & W. Schneider (eds.), *Diagnostik und Förderung des Leseverständnisses*, 19–44. Göttingen: Hogrefe.
- Schleppegrell, Mary J. 2004. *The language of schooling. A functional linguistic perspective*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Sirin, Selcuk. 2005. Socioeconomic status and academic achievement: a meta-analytic review of research. *Review of Educational Research* 75(3). 417–453.
- Skehan, Peter & Lucile Ducroquet. 1988. *A comparison of first and foreign language learning ability*. <http://www.opengrey.eu/item/display/10068/509355>.

- Sparks, Richard L., Jon Patton & Leonore Ganschow. 2012. Profiles of more and less successful L2 learners: A cluster analysis study. *Learning and Individual Differences* 22(4). 463–472.
- Udry, Isabelle, Amelia Lambelet & Raphael Berthele. 2019. *Language aptitude: Construct and empirical findings. Summary of literature review*. Tech. rep. Fribourg.
- Weiß, Rudolf H. 2006. *Manual Grundintelligenztest Skala 2 – Revision (CFT 20-R)*. Göttingen: Hogrefe.
- Wen, Zhisheng (Edward), Adriana Biedroń & Peter Skehan. 2017. Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching* 50(1). 1–31. DOI: 10.1017/S0261444816000276.

Chapter 10

The stability of language aptitude: Insights from a longitudinal study on young learners' language analytic abilities

Isabelle Udry^{a,b} & Jan Vanhove^a

^aUniversity of Fribourg, Institut de Plurilinguisme ^bZurich University of Teacher Education

An enduring question in aptitude research is the extent to which aptitude is a stable trait or a time-varying attribute. If aptitude were a perfectly stable trait, interindividual differences in aptitude at one point in time should be perfectly correlated with interindividual differences at a later point in time. However, raw test scores are affected by measurement error, a result of which is that correlations between raw test scores at different points in time underestimate the correlations between the actual skills measured by these tests at different points in time. The analyses of the longitudinal LAPS II aptitude data ($n = 636$; translated and adapted versions of MLAT and PLAB subtests) take into account measurement error and indicate that the children's ability to solve the MLAT and PLAB tests at the first data collection (autumn 2017, mean age: 10;5 years) and their ability at the third data collection (spring 2019, mean age: 12;1 years) are correlated at $\rho = 0.74$ (95% CrI: [0.69, 0.79]). This suggests that the ability to solve the two aptitude tests is not a perfectly stable interindividual trait, but that, by and large, interindividual differences are nonetheless maintained over the course of one-and-a-half years of cognitive development.

1 Introduction

Language aptitude as defined by Carroll (1958) consists of four basic components, i.e. phonetic coding ability, grammatical sensitivity, inductive ability, and rote



Isabelle Udry & Jan Vanhove. 2021. The stability of language aptitude: Insights from a longitudinal study on young learners' language analytic abilities. In Raphael Berthele & Isabelle Udry (eds.), *Individual differences in early instructed language learning: The role of language aptitude, cognition, and motivation*, 197–209. Berlin: Language Science Press. DOI: 10.5281/zenodo.

5464763 

memory (see Chapter 1 for a discussion). One of the key debates in aptitude research is whether the construct is a stable characteristic or an ability that can be developed. Addressing this issue contributes to establishing a conceptual aptitude framework. It could also clarify whether fostering aptitude components enhances language learning. Despite educational and theoretical relevance, studies dealing with the stability of language aptitude in general, and particularly in children, remain scarce.

Aptitude stability can be explored in the data in different ways: 1. Researchers consider average aptitude test scores achieved by one or several groups of participants (we refer to this as the *group averages* approach) or 2. they determine an individual's relative ranking within the group (which we call the *relative ranking* approach). In both approaches, researchers then look for patterns, either cross-sectionally (by comparing groups at different developmental stages) or longitudinally (by comparing data obtained at different times from the same individuals or groups). How these patterns are explained depends amongst other things on the researchers' conceptualization of construct stability: Developmental changes (expressed as age-related gain scores obtained in an aptitude test) can be interpreted as construct malleability. Alternatively, such changes can be seen as indicating construct stability if an individual's ranking within the population or group remains largely constant, despite increased aptitude scores at different times of testing.

In our view, describing changes in average scores only is scarcely insightful in our context, as children are expected to score higher on aptitude tests as they mature, namely due to developmental changes in cognition. It would be more revealing to find out if they progress at the same rate (indicating a general developmental pattern and a stable trait) or differentially (indicating individual developmental patterns and therefore a malleable ability). Depending on whether the *group averages* and *relative ranking* change, different conclusions may be drawn:

1. Group averages and the relative ranking both change: possible indication of malleability
2. Group averages change, the relative ranking remains the same: possible indication of stability
3. Group averages do not change, but the relative ranking does (due to some individuals progressing while others' ability to solve the test decreases): possible indication of malleability
4. Neither group averages nor relative rankings change: possible indication of stability

Several authors assume that language aptitude is subject to change (see for instance Grigorenko et al. 2000, Singleton 2017). Their view runs contrary to earlier conceptions, as expressed by Carroll (1981: 86) who described language aptitude as “relatively hard to modify in any significant way.” In the same article, the author (1981: 84) slightly qualified his statement by adding that the initial aptitude components could be modelled as “more or less enduring characteristics” and as a “current state”.

The view of a stable characteristic has been substantiated in particular with evidence from a long-term study by Skehan (1986) and Skehan & Ducroquet (1988) which found that some measures of L1 development, namely L1 vocabulary and mean length of utterance, were predictive of L2 aptitude measures assessed 13 years later in the same participants.

It is worth remembering that Carroll, who first conceptualized language aptitude, was mainly concerned with capturing a snapshot of people’s potential before they started learning a language in order to predict their later L2 achievement. Whether this potential was innate or malleable was not an explicit question in these early stages of aptitude research. However, as early as 1964, careful readers could detect Carroll’s awareness of the issue tucked away in a footnote, stating that the extent to which “the behaviour measured on the aptitude tests ... can be modified by training” would still “need to be investigated” (Carroll 1964: 89). Later, he expressed doubts about the feasibility of such a training, stating that “there are some general grounds for pessimism regarding the teaching of aptitudinal skills” (1973: 8), contributing to his view was the fact that not enough research had been conducted on the matter.

2 Review of the literature

Studies investigating the stability of language aptitude are rather scarce and usually underpinned by the hypothesis that aptitude is shaped by language experience. Roehr-Brackin & Tellier (2019) examined the relationship and development of language aptitude and metalinguistic awareness with 111 anglophone beginning learners of L2 French aged 8 to 9 years. In phase 1 (16 weeks) of the project participants were divided into four groups that were taught either German, Italian, Esperanto, or Esperanto with an added focus-on-form element. In phase 2 (16 weeks), all children learnt French with an element of focus-on-form. The authors administered tests for aptitude (an adaptation of the MLAT-E by Carroll & Sapon (1976) for British English speakers), metalinguistic awareness, and L2 French proficiency (listening, reading, writing, grammar) in a pretest–posttest

design which included immediate and delayed posttests for L2 French. The authors detected increases in aptitude test scores with a medium effect size. According to the “*group averages*” definition of stability given in the introduction, they concluded that language aptitude was dynamic in the sample. Moreover, the authors found that children who performed well on the aptitude pretest also did well on the aptitude posttest, and vice versa, suggesting a largely stable ranking. This finding was interpreted in favour of the aptitude test’s predictive value for young learners’ L2 performance. However, based on the “*relative ranking*” definition of stability, we argue that this finding could also point to the stability of language aptitude.

An increase in aptitude test scores with age was also detected by Suárez Vila-gran (2010) who administered Spanish and Catalan translations of the MLAT-E (see §2.2) to 629 Spanish-Catalan bilingual learners of English aged 8 to 15. She observed a considerable increase in aptitude scores between ages 8 and 9. After the age of ten, gain scores weren’t as large, and the author therefore suggests that aptitude may stabilize around age 11.

Kiss & Nikolov (2005) tested aptitude (with a Hungarian MLAT based test), motivation, and English proficiency (listening, reading, writing) in 419 12-year-old L2 English learners. The authors also recorded time of exposure to English at school and in private tuition which ranged considerably from 100 to 1085 hours ($M = 343$; $SD = 131$). Kiss and Nikolov explored the effects of language experience on aptitude by establishing correlations between time spent on learning and aptitude test scores. This correlation being weak, they concluded that language aptitude in the Carrollian sense did not improve with “the amount of time used for practice and exposure” (Kiss & Nikolov 2005: 134).

In a subsequent study, Kiss (2009) reconsidered the interplay between language experience and aptitude. The author administered a Hungarian aptitude test for young learners to 52 Hungarian children. The aim was to select 26 8-year-olds for a newly established bilingual English-Hungarian teaching program. The author compared the results from all 8-year-old children to those from 12-year-olds from a previous study. She found that the 12-year-olds performed better on the vocabulary learning subtest than their younger counterparts. Kiss (2009: 268f) speculates that these differences are owed to the older children’s greater language learning experience and knowledge of strategy use. Referring to the idea that increased group averages reflect aptitude malleability, the author argues that language aptitude is dynamic, at least up to the age of 12.

A frequently cited study by Sáfár & Kormos (2008) addressed the stability of the construct with 61 Hungarian learners of English aged 15 to 16 years. The authors assessed language aptitude and short-term memory at the beginning

and end of the academic year. 41 participants followed an English-Hungarian bilingual program (with sixteen 45 min English lessons per week + 4 × 45 min. CLIL per week) and 21 participants were from a regular Hungarian secondary school (with 4 × 45 English lessons a week predominantly communicative with some focus-on-form instruction). Aptitude test scores increased significantly in both groups between the two data collections, independent of the intensity of instruction. Learners in the bilingual program, however, improved more than their counterparts in the regular setting. Based on these findings, the authors concluded that language aptitude is dynamic and changes with language experience. The authors also stated that language aptitude appears to be less relevant in communicative teaching with a focus-on-form element. It seems worth noting that some information on other potential influences, such as general learning abilities or admission criteria for the bilingual program, may have contributed to explaining the results.

In summary, current empirical findings suggest that language aptitude is dynamic in children younger than 12. Primary school children score higher on aptitude tests as they get older, with important increases in test scores being observed between 8 and 9 years (Milton & Alexiou 2006, Kiss 2009, Suárez Vilagran & Muñoz 2011). These findings are based on the *group averages* definition of stability, i.e. a gradual improvement in the average performance of groups of children.

3 Method

We investigated the stability of language aptitude in primary school children over a period of 1.5 years. To this aim, we defined language aptitude as language analysis (Skehan 1998), i.e. the grammatical sensitivity and inductive ability components from Carroll's (1958) construct definition (see also Chapter 1 for a discussion). As outlined earlier, primary school children are still maturing cognitively, and their aptitude scores are expected to improve with age. Therefore, we were more interested in the extent to which individual differences in language aptitude remain stable over time, i.e. whether participants who perform well relative to other participants on an aptitude test at one point T1 will still do so at a later time of testing T2 (or even T3, as in our data). This can be inferred from the correlation between the participants' test results at T1 and T2 (and T3): The stronger this correlation is, the more stable interindividual differences in language aptitude are.

4 Participants and procedure

The study design is fully described in Chapter 2 and will be summarised briefly: To assess language analytic ability (i.e. grammatical sensitivity and inductive ability), we adapted the following tests for German-speaking young learners¹: MLAT-E subpart “Matching Words” on grammatical sensitivity (Carroll & Sapon 1976) and PLAB subpart 4 on inductive ability (Pimsleur et al. 2004). The same participants from LAPS II (4th and 5th graders at T1), completed these tests at three different times T1–T3: T1 = Autumn 2017 (mean age 10;5); T2 = Spring 2018 (mean age 11); T3 = Spring 2019 (mean age 12;1). A total of 636 participants completed the tests between T1 and T3. Figures 1 and 2 show the correlations between the test scores from T1 to T3. Table 1 summarizes the test results.

Table 1: MLAT and PLAB test results at T1–T3. The total number of participants completing the tests at different times between T1–T3 is 636. MLAT has 30 items and PLAB 15.

Test	Time	Participants	Mean ^a	SD
MLAT	T1	590	0.48	0.23
	T2	575	0.58	0.22
	T3	565	0.66	0.19
PLAB	T1	596	0.37	0.19
	T2	577	0.44	0.22
	T3	564	0.56	0.25

^aProportion of correct answers

Clearly, the results from the three testing times do not correlate perfectly (this would be indicated by a correlation coefficient of +1). Even if the ability to solve the MLAT or the PLAB were interindividually stable, we would still not expect to see a perfect correlation. This is due to measurement error: If two latent variables that correlate perfectly with each other are measured imperfectly, these measurements will still not yield a perfect correlation.

If we knew the measurement error or the reliability coefficient of the tests, we could solve this problem by calculating disattenuated correlations. While we do not know the actual reliability of the instruments, we can estimate them. The reliability coefficients (ω_{RT} , McNeish 2018, Revelle 2019) for the MLAT variable are

¹We would like to thank Charles W. Stansfield for permission to translate and adapt parts of the MLAT-E and PLAB for this study.

0.90 (T1), 0.89 (T2) and 0.88 (T3); for PLAB: 0.74 (T1), 0.80 (T2) and 0.84 (T3). If we use these values to disattenuate the correlations (using the `correct.cor()` function in the `psych` package for R, William 2018), we obtain the values in Table 2.

Table 2: Disattenuated correlations MLAT and PLAB

	MLAT		PLAB	
	T1	T2	T1	T2
T2	0.77		0.60	
T3	0.62	0.78	0.61	0.65

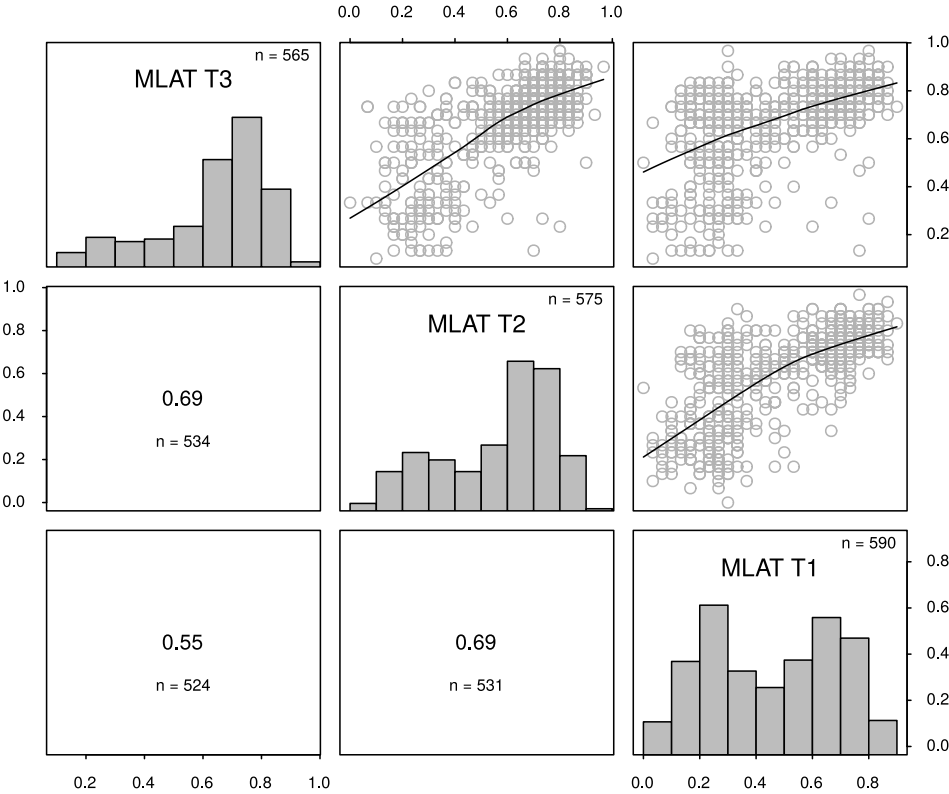


Figure 1: Scatterplot matrices with the MLAT results at the three data collections. Upper triangle: Scatterplots with scatterplot smoothers. Main diagonal: Histograms. Lower triangle: Pearson correlation coefficients as well as the number of data points on which these were based.

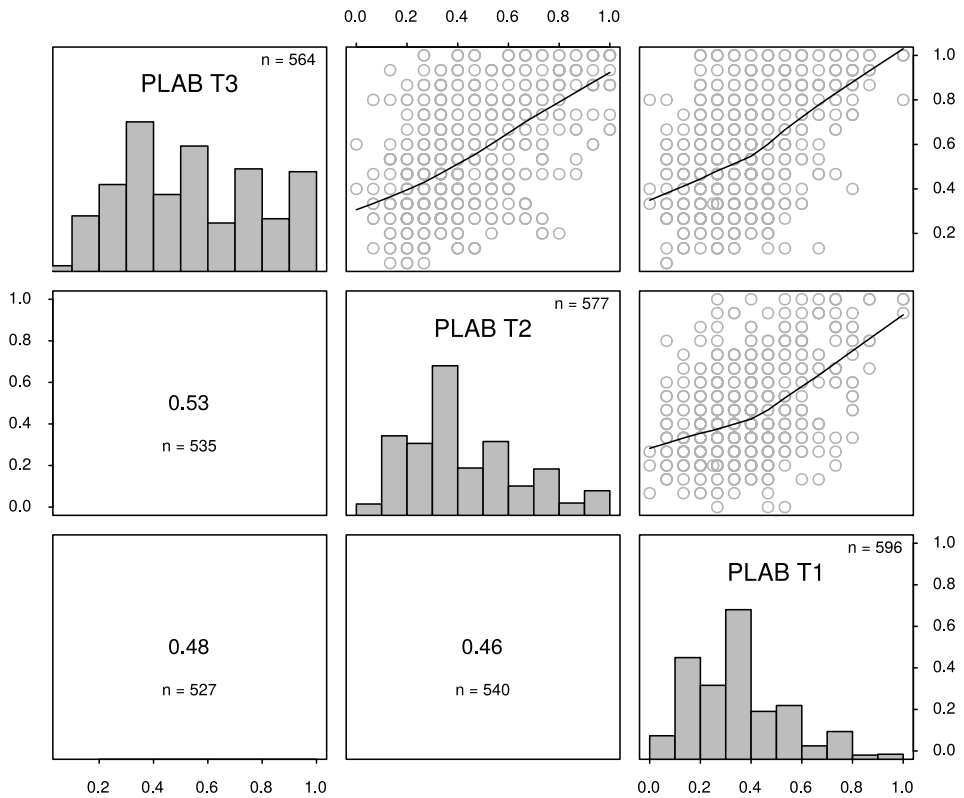


Figure 2: PLAB results and correlations at different times of testing T1–T3. Scatterplot matrices with the MLAT results at the three data collections. Upper triangle: Scatterplots with scatterplot smoothers. Main diagonal: Histograms. Lower triangle: Pearson correlation coefficients as well as the number of data points on which these were based.

While computing disattenuated correlation coefficients is fairly straightforward, this analysis does not consider the dependence that exists between the tests: The MLAT scores for T1, T2 and T3 are based on the same test items, and the same goes for the PLAB scores. To take these dependencies into account, we ran an alternative, if more involved, analysis in which the participants' item-level responses were modelled. This analysis was run using generalized (logistic) mixed-effects models, which are capable of estimating the participants' latent abilities as well as the items' latent difficulties.

In language research, analyses based on mixed-effects models usually focus on the fixed effects, but in our case, it is the random effects that are of particular interest. For each participant, the latent ability to solve the MLAT and PLAB can

be estimated for T1, T2 and T3. Also, the correlations between the participants' estimated latent abilities at T1, T2 and T3 can be estimated. In doing so, the analysis can also take into account the fact that items vary in their difficulty and that the relative difficulty of the test items may vary between T1, T2 and T3.

We fitted three models: one on the MLAT responses, one on the PLAB responses and one on all responses combined. The models we fitted were Bayesian generalized (logistic) mixed-effects models; for our purposes, Bayesian models have the advantage that they can not only estimate the correlation between the participants' latent abilities at T1, T2 and T3, but also quantify the uncertainty about this estimation. The models were fitted using the `brm()` function in the `brms` package for R (Bürkner 2017). "Result" is a binary variable that indicates for each individual response whether it was correct (1) or not (0). "Time" is a categorical variable with three levels (T1, T2, T3), "Item" is a categorical variable specifying the test item, and "StudentID" is a categorical variable specifying the participant. The three models were specified as follows (in `brms` notation):

```
m <- brm(result ~ 0 + Time + (0 + Time | Item) + (0 + Time |
  StudentID), data = d, family = bernoulli(link = "logit"), cores
  = 4, iter = 4000, warmup = 1000)
```

This model estimates (a) the probability (in logits) of a response being correct at Times 1, 2 and 3, (b) between-item differences in this probability, (c) between-participant differences in this probability, (d) the correlations among the between-item differences at Times 1, 2 and 3, and (e) the correlations among the between-participant difference at Times 1, 2 and 3. For our purposes, (e) is what is important, viz., the extent to which differences among the participants' abilities at Time 1 are maintained at Times 2 and 3.

5 Results

The MLAT model was fitted on 51,875 responses (30 items \times 636 participants \times 3 data collections, with some missing data); the PLAB model was fitted on 26,044 responses (15 items \times 636 participants \times 3 data collections, with some missing data); the combined model was fitted on 77,919 responses (45 items \times 636 participants \times 3 data collections, with some missing data).

Tables 3–5 summarise the main results pertinent to our research question. The full model output can be inspected at <https://osf.io/pf5g8/>. Posterior predictive checks indicated that the models reported here can generate the key characteristics of the dataset; these checks are also reported in the online appendix.

Overall, correlations range from moderate to strong (0.63–0.83). The MLAT correlations are stronger (0.65–0.79) than the PLAB correlations (0.63–0.68) and even stronger correlations are obtained when the two tests were considered together (0.74–0.83). Also, correlations are stronger for short intervals (T1–T2; T2–T3) than for the longest period T1–T3.

The first model estimates that the differences in solving the MLAT (expressed in logits) is correlated between T1 and T2 at 0.78 (95% uncertainty interval: [0.74; 0.83], between T2 and T3 at 0.79 ([0.74; 0.84]), and between T1 and T3 at 0.64 ([0.58; 0.71]).

For the PLAB, this ability was estimated to correlate with 0.63 (95% uncertainty interval: [0.54; 0.72]) from T1 to T2, 0.68 from T2 to T3 and 0.66 between T1 and T3.

In a third model, the overall ability to solve both MLAT and PLAB was analysed in the same way. Results show that language analytic abilities as measured by both tests were correlated at 0.83 ([0.78 0.86]) between T1 and T2, 0.82 ([0.78 0.86]) between T2 and T3 and 0.74 ([0.69 0.79]) from T1 to T3.

6 Discussion

We were interested in the stability or relative development of language-analytic ability as a subcomponent of language aptitude. Language-analytic ability was assessed in 636 primary school children aged 10–12 years at three times over 1.5 years with adaptations of the MLAT-E Matching Words (grammatical sensitivity) and the PLAB subtest for inductive ability. Our findings indicate that overall results improved over time, with test scores being highest at T3 for both MLAT and PLAB. A gradual increase in test scores was expected due to the children's cognitive maturation. As one reviewer pointed out, higher scores may also be linked to test familiarity. Even though the MLAT and PLAB were administered at intervals of 6 months (T1–T2) and 12 months (T2–T3), practice effects cannot be ruled out entirely and it is possible that maturation and test familiarity are intertwined to some degree.

We also adopted an interindividual perspective, assessing whether our participants' *relative* ability to solve the aptitude tests remained stable with increasing age. Recall that by latent ability or relative ability to solve the aptitude tests, we mean the correlation of the *relative difference* of the test scores between testing times. In other words, this correlation indicates how strongly the ranking among participants based on their scores has changed over time.

We will discuss the longest interval T1 to T3 which most adequately mirrors long-term changes in our data. Correlations between T1 and T3 are strong

Table 3: The estimated correlations between the relative differences of participants' abilities to solve the MLAT test (expressed in logits) and their 95% credible intervals.

Parameter	Est.	95% credible interval	
		lower bound	upper bound
\hat{P} T1-T2	0.78	0.74	0.83
\hat{P} T1-T3	0.65	0.58	0.71
\hat{P} T2-T3	0.79	0.74	0.84

Table 4: The estimated correlations between the relative differences of participants' abilities to solve the PLAB test (expressed in logits) and their 95% credible intervals.

Parameter	Est.	95% credible interval	
		lower bound	upper bound
\hat{P} T1-T2	0.63	0.54	0.72
\hat{P} T1-T3	0.66	0.57	0.74
\hat{P} T2-T3	0.68	0.61	0.75

Table 5: The estimated correlations between the relative differences of participants' abilities to solve the MLAT and PLAB tests (expressed in logits) and their 95% credible intervals.

Parameter	Est.	95% credible interval	
		lower bound	upper bound
\hat{P} T1-T2	0.83	0.78	0.86
\hat{P} T1-T3	0.74	0.69	0.79
\hat{P} T2-T3	0.82	0.78	0.86

($\rho = 0.74$) when both tests are considered together, and moderate when the tests are considered separately, i.e. $\rho = 0.65$ for the MLAT and $\rho = 0.66$ for the PLAB. These results suggest that language-analytic ability, as operationalised by these tests, is not entirely stable, as this would have been evidenced by correlations in an even higher range. At the same time, moderate to strong correlations indicate that a relationship between language-analytic ability at T1 and T3 is still present in the data. In conclusion, the ability to solve the two aptitude tests is not a perfectly stable interindividual trait, but, by and large, interindividual differences were nonetheless maintained over the course of 1.5 years of cognitive development.

References

- Bürkner, Paul-Christian. 2017. Brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80(1). 1–28.
- Carroll, John B. 1958. A factor analysis of two foreign language aptitude batteries. *The Journal of General Psychology* 58(1). 3–19. DOI: 10.1080/00221309.1958.9710168.
- Carroll, John B. 1964. The prediction of success in intensive foreign language training. In *Training research and education*, 87–136. Pittsburgh: University of Pittsburgh Press.
- Carroll, John B. 1981. Twenty-five years of research on foreign language aptitude. In Karl C. Diller (ed.), *Individual differences and universals in language learning aptitude*, 83–118. Rowley: Newbury House.
- Carroll, John B. & Stanley M. Sapon. 1976. *Modern language aptitude test – Elementary*.
- Grigorenko, Elena L., Robert J. Sternberg & Madeline E. Ehrman. 2000. A theory-based approach to the measurement of foreign language learning ability: The Canal-F theory and test. *The Modern Language Journal* 84(3). 390–405.
- Kiss, Csilla. 2009. The role of aptitude in young learners' foreign language learning. In Marianne Nikolov (ed.), *The age factor and early language learning*, 253–276. Berlin: De Gruyter Mouton.
- Kiss, Csilla & Marianne Nikolov. 2005. Developing, piloting, and validating an instrument to measure young learners' aptitude. *Language Learning* 55(1). 99–150.
- McNeish, Daniel. 2018. Thanks coefficient alpha, we'll take it from here. *Psychological Methods* 23(3). 412–433. DOI: 10.1037/met0000144.

- Milton, James & Thomai Alexiou. 2006. Language aptitude development in young learners. In C. Abello-Contesse, R. Chacón-Beltrán & M. D. López-Chiménez (eds.), *Age in L2: Acquisition and teaching*, 177–192. New York: Peter Lang.
- Pimsleur, Paul, Daniel J. Reed & Charles W. Stansfield. 2004. *Pimsleur language aptitude battery PLAB*. Rockville: Second Language Testing.
- Revelle, William. 2019. *Using R and the psych package to find ω* . <http://personality-project.org/r/psych/HowTo/omega.pdf>.
- Roehr-Brackin, Karen & Angela Tellier. 2019. The role of language-analytic ability in children's instructed second language learning. *Studies in Second Language Acquisition* 41. 1111–1131.
- Sáfar, Anna & Judit Kormos. 2008. Revisiting problems with foreign language aptitude. *International Review of Applied Linguistics in Language Teaching* 46(2). 113–136.
- Singleton, David. 2017. Language aptitude: Desirable trait or acquirable attribute? *Studies in Second Language Learning and Teaching* 7(1). 89–103.
- Skehan, Peter. 1986. Cluster analysis and the identification of learner types. In Vivian J. Cook (ed.), *Experimental approaches to second language acquisition*, 81–94. Oxford: Pergamon Press.
- Skehan, Peter. 1998. *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, Peter & Lucile Ducroquet. 1988. *A comparison of first and foreign language learning ability*. <http://www.opengrey.eu/item/display/10068/509355>.
- Suárez Vilagran, Maria Del Mar. 2010. *Language aptitude in young learners: The elementary modern language aptitude test in Spanish and Catalan*. Barcelona: Universitat de Barcelona. (Doctoral Dissertation).
- Suárez Vilagran, Maria Del Mar & Carmen Muñoz. 2011. Aptitude, age and cognitive development: The MLAT-E in Spanish and Catalan. *EUROSLA Yearbook* 11(1). 5–29. DOI: 10.1075/eurosla.11.03sua.
- William, Revelle. 2018. *Psych: Procedures for personality and psychological research*. <https://www.scholars.northwestern.edu/en/publications/psych-procedures-for-personality-and-psychological-research>.

Chapter 11

Summing up: Individual differences in primary school foreign language learning

Raphael Berthele^a & Isabelle Udry^{a,b}

^aUniversity of Fribourg, Institut de Plurilinguisme ^bZurich University of Teacher Education

In Chapter 1, we set the stage for this volume with an outline of the constructs, research methods, and pedagogical relevance of individual differences (IDs) in foreign language teaching and learning. At the close, we wish to revisit the main findings of the chapters of this volume in order to summarize what we have learnt from the investigation and to think about what our insights can mean both for researchers and practitioners.

Our ambition was to provide empirical evidence on learner characteristics that explain (statistically and theoretically) variance in second language skills. The long-term goal of such endeavours is to come to a better understanding of learner variability which in the future could inform pedagogical choices and practices in the foreign language classroom.

In the following, we summarize the main findings drawn from the different chapters and discuss their educational and theoretical implications. The road leading from research as presented in this volume to policy recommendations is long, winding, and often rocky. Language curricula and multilingual education, eminently so in officially multilingual countries, are never entirely “evidence-based”, i.e. they are not simple transpositions of research findings into practice. They are the result of complicated political and other institutional processes, with fluctuating and often inconsistent recourse to scholarly research (see Berthele



2019 for examples). Our goal, however, is to contribute to the growing pool of evidence on the possibilities and the limits of foreign language learning in compulsory primary school curricula.

1 Summary of the findings

In our research, we made an attempt to contribute to three different perspectives on individual differences (IDs) in learning foreign languages: The first is the interest in a better understanding of the variables or constructs that account for the differences in the ability to learn a foreign language: What are the cognitive, affective, and sociological variables that are related to foreign language learning ability, and what is the internal dimensionality of such a broad array of individual difference measures? The second perspective relates to the feasibility, based on the results produced within the first perspective, of drawing on these ID variables in order to prognosticate language development in the foreign language. The third perspective is the interest in change or stability of ID variables over time.

To contribute new evidence to all three perspectives, we examined a range of ID variables deemed to be associated with foreign language learning at primary school beyond the components theorized by Carroll in the 1950s and 60s.

1.1 Dimensions

Of particular interest to the field is the relationship between cognitive ID variables that are more (phonetic coding ability, grammatical sensitivity, inductive ability) or less language related (e.g. fluid intelligence, working memory, field independence). Exploratory and confirmatory factor analyses on two independent samples (LAPS I and II) showed that general and language related cognitive variables load on the same factor. We chose to call this factor *Cognition/Aptitude*. Since the structure of constructs that emerged from the exploratory factor analyses was confirmed in the second, larger sample, we confidently concluded that for the age group of 10- to 12-year-old children examined here, general cognitive abilities and language-oriented abilities represent a single dimension (cf. Chapter 3). In line with other evidence on L2 motivation, we identified a distinction among variables that can broadly be delimited into intrinsic vs. extrinsic facets of motivation. Our analyses yielded two affective factors we named *L2 Academic Emotion* and *Extrinsic* factor. A regression analysis showed that the *L2 Academic Emotion* factor, together with *Cognition/Aptitude*, relates positively

to L2 proficiency. The *Extrinsic* factor, when considered together with the other two factors, was associated negatively with L2 proficiency.

Further explorations of these data are presented in Chapters 5 and 9. In Chapter 9, the analyses compare the linear relationships of individual difference variables with the language of instruction (German) and the foreign language (English), respectively. The analyses yield similar patterns of associations of these individual characteristics with both languages, a result in line with the assumption of an underlying ability to learn and use languages that is not dedicated to a specific language. Moreover, Chapter 9 also explores the contribution of socioeconomic variables and home language use (i.e. individual multilingualism beyond the languages taught at school). These associations are further explored in Chapter 5, by distinguishing between cultural and economic dispositions of the children. The analyses presented in Chapter 5 suggest that such family background characteristics are indeed related to foreign language skills in the target language English. Their contribution, however, is indirect via the two factors that stood out in Chapter 3, namely *Cognition/Aptitude* and *L2 Academic Emotion*. The estimate of the direct path to English as a foreign language is small. The chapter also investigates whether specific features of pupils particularly associated with socio-educational vulnerability, such as being born abroad or not speaking the local language German at home, impact English skills beyond the factorial structure already identified in Chapter 3. These analyses show that the variable cultural and economic dispositions are first and foremost associated with the two factors emerging in Chapter 3 but do not affect L2 English skills beyond these indirect effects.

1.2 Prediction

Identifying predictor variables for L2 proficiency is a way to assess student potential (Chapter 4). Assessing this potential had been the goal of the early language aptitude tests (see Chapter 1 for details). Our study focused on a markedly different educational context than the early and classical aptitude tests did: Two foreign languages are an obligatory part of the curriculum, and the learners are not young adults or adolescents, but children in primary school.

We first trained statistical models with all variables available from the test battery on the LAPS II training set. That is, motivational, general and language related cognitive variables, as well as social background variables were all taken into consideration when training these models. This yielded a *No costs spared model* whose power to prognosticate English skills at T3 was assessed on an unseen test set. Next, we compared this model to simpler, more practical models

with a small set of measures that are relatively easy to take in a classroom. We refer to these models as *Cheap models*. Moreover, since our participants already had English skills, we were also interested in how the English test at T1 fared as a sole predictor for English proficiency at T3.

Overall, our results showed that the English test taken at T1 provides a good forecast of L2 proficiency 1,5 years later at T3. When all variables were considered, i.e. the *No costs spared* model, a linear regression model with 7 predictors fared best: Variables included in this *No costs spared* model are English T1, grade (=year in school), intrinsic motivation, self-concept English, L1 German, MLAT (grammatical sensitivity), and PLAB (inductive ability). The practical, or *Cheap*, models contain small selections of variables that encode either information on learners that is usually readily available in a classroom setting or that is rather easy to collect (e.g. motivational information via questionnaires). The best of these cheap models is one that includes the English test scores at T1 and such easily obtained information. This model performed only slightly worse than the 7 predictor *No costs spared* model at T1. Overall, the comparison of these different predictive models shows that the most informationally rich single measurement for prognostication of skills at T3 is a test of the same skill at T1. A high level of stability within constructs is also what emerges from the analysis in Chapter 10. Here, the stability of two language related measures, a grammatical sensitivity task inspired by an MLAT form and an inductive ability task inspired by a PLAB form are investigated longitudinally: When accounting for measurement error, the scores between T1 and T3 are not perfectly correlated. Thus, the abilities measured are not perfectly stable traits within learners across time. However, the high association ($\rho = 0.74$) suggests a high level of stability of these measurements.

When interpreting the prognostic models in Chapter 4, it is important to keep in mind that variables or dimensions that have not emerged as predictors in these models must not be hastily dismissed as being irrelevant for foreign language learning. Rather, our results suggest that considered together, the variables from the *No costs spared* model give, with some precision, an estimate of how a student's L2 skills are likely to develop between T1 and T3, i.e. between grades 4 and 6. The *English only model* and some of the *Cheap models* fared slightly worse in doing so. As can be seen from the results reported in Chapter 4, the mean error of the models differed within a span of 1.9–2.2 points on a 20-point scale when fitted on the test set. We are not aware of other studies that have attempted to assess in this fashion if and how accurately prognostications about foreign language learning can be made. The main insight emerging from our analyses is

that prognostic testing in primary school foreign language learning is indeed possible, and it can be done with high predictive accuracy.

1.3 Motivation and creativity

Our investigation did not only cover abilities (language related or general cognitive), but also affective, attitudinal and motivational dispositions, as called for by scholars for quite some time (Dörnyei 2010: 267, Parry & Stansfield 1990: 2). Beyond the insights into the different factorial dimensions provided in Chapter 3, two chapters are dedicated to a more fine-grained analysis of different motivational components as well as their development over time (Chapters 7 and 8). The specific setting of our learners, hailing from two different areas of the German-speaking part of officially quadrilingual Switzerland, called for a differentiated analysis of the pupils' motivational stances regarding their own foreign language learning. The comparison of two foreign languages (English and French), both compulsory but introduced in the curriculum in reverse order in the two settings, reveals insights both into the effects of spatial proximity of the target language territory and of foreign language subjects in the curriculum. French is the demographically and culturally dominant language in the region of LAPS I and it is taught as the first L2 in this region's German-speaking areas. The proximity of a French speech community was expected to trigger higher motivation in these pupils, at least with respect to extrinsic or *lingua franca* uses of the language. The results reported in Chapter 7 show that there is no difference between the two areas, that is the motivational dispositions look the same as in the region of LAPS II where French is taught as the second foreign language (L3) and is much further away, both geographically and culturally. The motivation to learn English is higher across all motivational sub-dimensions and in both of the contexts investigated.

Moreover, as in the other developmental analyses, the development across time of the constructs seems rather stable. The most consistent change, as discussed in Chapter 8, seems to be the decline in school-related motivation for both target languages across the 1,5 years investigated. The decline is not dramatic and similar tendencies have been observed in numerous studies on different subjects (see Shan 2020, chapter 2.1 for discussion and references).

Motivation and creativity, more specifically divergent thinking, are investigated in Chapter 6. It is conceivable that task-based learning that requires pupils to generate creative outcomes such as poems or role plays, draws particularly on creative thinking. In such an environment, creative thinking is expected to be associated with the motivation to learn the target language(s). However, as the

analyses reveal, no noteworthy association of the two constructs can be found. Modelling the association of creativity with L2 proficiency in both languages (French and English) in the LAPS I sample reveals a weak but significant association. As discussed in Chapter 6, the causality or contribution of one construct on the other remains a matter of speculation and is difficult to pin down.

2 Discussion of the main findings

2.1 Applying aptitude-related findings

Aptitude tests were initially developed in the US to select able students for language classes (see Chapter 1 for a discussion). In new educational paradigms, the role of predictive testing as a tool of selection seems less fitting. Public education is expected to do its best to even out achievement gaps to which individual differences contribute, and to accommodate different learner needs. At the same time, selection is one of the functions of modern educational systems. Overt or covert selection practices are therefore in place in contemporary educational contexts, also in the Swiss case our study was concerned with. Foreign language learning being part of the compulsory subjects in the Swiss curricula, achievement in these subjects must therefore be considered when educational selection processes are at stake. Achievement measures in foreign language subjects sometimes form a part of high-stakes test results feeding into tracking decisions. In this context, we feel that we can nourish current debates that concern such language-related selection processes. Individual differences are at the centre of attention when discussing whether some students should be exempted from one or both compulsory foreign language subjects if they suffer from learning difficulties (cf. the Linguistic Coding Difference Hypothesis mentioned in Chapter 1). If dispensation from, say, English as a foreign language were to be considered a reasonable practice by teachers and other stakeholders, it would indeed be possible to draw on the results presented here – in the sense that metrics that are associated with L2 skills as discussed in Chapters 3 and 4 could be used to define and assess selection criteria. However, as mentioned in Chapter 1, identifying individual differences in predispositions for learning foreign languages can serve other purposes as well. It can be used, for example, to recognize learners who need more time to acquire similar skills.

In Chapter 1, we critically discussed the aptitude-treatment-interaction (ATI) approach which investigates the mutual influence between language aptitude and teaching methods. ATI is underpinned by the assumption that a) learners do have different aptitude profiles and b) that these profiles are in interaction

with teaching methods, i.e. selecting a method tailored to the aptitude profile will enhance learning. Due to various reasons, ATI has so far failed to deliver enough robust evidence to substantiate these claims. While it is uncontroversial that individual differences in learners exist, and also that learners express different preferences when it comes to choosing ways and methods of learning, the claim that differentiated pedagogical treatments allow learners to systematically draw on strengths and compensate for weaknesses is not based on sufficient robust evidence (see our discussion in Chapter 1). If an ATI approach nevertheless should turn out to be empirically sound, the dimensionality of the aptitude construct from Chapter 3 provides insights on the axes along which learners vary.

If one wishes to assess students' potential L2 development, any of the prognostic models discussed in Chapter 4 are of some avail. Their usability depends on context and purpose. If learners already have basic L2 knowledge, an English test is a suitable and convenient choice for teachers. However, the *Cheap models* are informative if one wishes to gain some insight into affective and cognitive dimensions, an area that is likely to be of interest to practitioners over and above the prediction of L2 skills. Intervention-based research could clarify whether such easy to collect information from the Cheap models could be used to adapt the pedagogical setting in order to attain better learning results longitudinally. If a Cheap model indicates, for instance, that a pupil's L2 motivation is low, appropriate measures could be taken to assist that student.

The extent to which education can influence learner performance at all may not always be as large as educators would like it to be. As shown in Chapter 5, socioeconomic status, which cannot easily be changed by the individual, bears strongly on the two factors that are positively but indirectly associated with L2 proficiency (via the constructs *Cognition/Aptitude* and *L2 Academic Emotion*). If the assumption is that social dispositions contribute causally to one or both of these constructs (and not vice versa), then this points to important hurdles for teachers and schools to change individuals' dispositions with respect to these two important constructs. This result raises concerns about how well an education system whose pledge is equal opportunity can live up to such expectations in real life. In a related vein, if the cognitive and/or linguistic abilities are partially predetermined by genetics, as suggested by Plomin (2019) or Stromswold (2001), this also points to limits of the extent to which individual differences can be pedagogically levelled out, in particular within the restricted possibilities of a dense curriculum in a state school with only limited time at disposal for L2 instruction.

2.2 Theory development

Inquiries into language aptitude usually frame the construct drawing on the Carrollian components, i.e. memory (associative or working memory), phonetic coding ability, grammatical sensitivity, and inductive ability – the latter two being also labelled *language analysis* in Skehan's (1998, see Chapter 2.1) three-component model. Based on the description of learner profiles by Skehan, many researchers distinguish between memory and analysis-oriented adult or adolescent learners.

For primary school children as examined in the LAPS project, a memory vs. language analysis distinction cannot be substantiated: As reported in Chapter 3, all general and language related cognitive variables load on one factor. Here, it is important to emphasize that in a study such as ours that encompasses a multitude of cognitive and affective variables, the measures of the constructs are unavoidably coarse: Digit span-based measures of working memory, for example, allow us to grossly distinguish between categories of students (they reliably identify students with serious working memory problems), but the measure, although widely used, is by no means the state-of-the art test of working memory capacity that one would apply if much more time were at one's disposal (e.g. operation span; cf. Conway et al. 2005 for a comparison and overview). It is important to resist the temptation to essentialize factors emerging from data as representing "scientifically proven reality", as more advanced tests might produce evidence pointing towards a more complex internal dimensionality of our component. Given the temporal and logistic constraints that also govern research as ours, the tests and measures we were able to generate represent what was feasible, not what would be done in an ideal world ruled by researchers. On the other hand, given the converging results of the exploratory and confirmatory factor analyses reported in Chapter 3, we feel confident affirming that for primary school learners aged 10–12, there is solid empirical evidence to postulate that memory and analysis related cognitive variables are highly positively correlated.

As discussed in Chapters 3 and 4, affective variables (particularly those connected to L2 Academic Emotion, i.e. intrinsic motivation, self-concept, and anxiety) make a separate contribution to L2 achievement in addition to cognitive variables. Based on these observations, we argue that future studies, be they within an ATI framework or not, could be enriched by going beyond the cognitive language-related focus of the Carrollian concept to include affective variables along the lines of self-determination theory (Deci & Ryan 2002), as a function of different teaching methods. In this view, assessing the effects of specific

treatment conditions on learner's self-concept, anxiety and pleasure in learning foreign languages seems – unsurprisingly – useful in the search for more efficient foreign language teaching, since our data show that these constructs are linearly associated with skills in the target language.

Focusing on the predictive value of the tests used in the study, aptitude measures for language analysis (grammatical sensitivity and inductive ability) have emerged as being predictive of achievement on several accounts: Both appear in the *No costs spared* model. Moreover, an inductive ability task is part of one of the *Cheap* models (Chapter 4) and a grammatical sensitivity task turns out to be a variable that predicts L2 English and L1 German proficiency as evidenced by the analysis in Chapter 9. On the one hand, these findings emphasize the importance of the factor *Cognition/aptitude* for foreign language learning. On the other, they suggest that within this factor, language-related tests are more strongly associated with L2 outcomes than measures of general intelligence or working memory. This is not surprising, since aptitude tests tap into language-oriented constructs, and language skills were the main outcome variables in our study. Based on our findings from the factor analysis, this indicates that the abilities required to solve tests for language analysis and tests for intelligence and WM overlap, with the language analysis tasks tapping into grammatical sensitivity and induction being more strongly associated with language-related outcomes.

In tackling individual predispositions to language learning comprehensively, we wanted to widen the scope by including potentially relevant, but lesser researched ID variables. Creativity has been hypothesized to play a role in foreign language learning, especially in the task-based approach adopted in Switzerland that relies on the learners' potential to generate ideas that help them solve communicative tasks using their target language skills. In multilingualism research, we find rather strong causal claims based on (weak) positive associations of creativity measures and the individuals' language repertoires (e.g. Fürst & Grin 2018). Our data suggest also that there is a positive linear association of non-verbal creativity and proficiency in foreign languages in the LAPS I test battery (see Chapter 6). The association is weak, though, and the direction of causal links cannot be established based on cross-sectional multivariate analyses. An association of creativity and motivation, however, could not be substantiated.

As we have discussed in Chapter 1, there are several unresolved issues in aptitude research, some of which can be elucidated by our findings. Language aptitude, as conceptualized by Carroll, has been deemed obsolete or at least out-dated by some. For example, in the wake of the *Natural approach*, scholars claimed that language aptitude is irrelevant for learning languages in such communicative, “natural” paradigms, in particular in the case of children (Skehan 2002: 72).

Notwithstanding such claims, the results in the literature reported in Chapter 1 as well as our findings show that “old-fashioned” language related constructs are clearly associated with the outcomes of language learning also in communicative language teaching/learning settings (Chapters 3, 9). The evidence also suggests that such aptitude measures contribute to forecasting learning outcomes in communicative teaching (Chapter 4).

Another matter of debate is the extent to which language aptitude is malleable or a relatively stable trait of the individual. This question is relevant if one considers training aptitude subcomponents, such as metalinguistic skills or language analysis, in order to improve learning outcomes. Testing such an assumption would require experimental settings. Our investigation cannot provide answers beyond the aforementioned relative stability of the MLAT words in sentences and PLAB inductive ability tasks over time (Chapter 10).

2.3 Future perspectives

To date, few studies have dealt with instructed foreign language learning at primary school on a large scale. Throughout this book we have explained our methodological and analytical choices, and we have also directed the readers’ attention to possible improvements to our research plan. We hope these discussions encourage other groups to build upon our work, benefit from our insights, and continue the investigation of individual differences in young language learners. Moreover, we would welcome experimental studies that extend on our results. These could clarify some of the theoretical and educational assumptions currently held about individual differences which are not sufficiently underpinned by empirical evidence. For example, we think that the claims made by ATI and its applicability should be examined further with carefully planned designs along the lines outlined throughout the book. Moreover, we have postulated the ability to solve language analysis tests (as in our grammatical sensitivity and inductive ability tasks) being rather stable within the individual when learning conditions are the same for all participants (i.e. they all followed the same curriculum). This does not license claims about the effects of a specific aptitude or metalinguistic training on proficiency. An experimental design with treatment and control groups would be insightful in this respect.

As discussed in Chapters 3, 7 and 8, motivational dimensions of foreign language learning are undoubtedly an important aspect of individual variability in language learning. Our results corroborate previous findings that children are generally keen to learn foreign languages at school and that their motivational dispositions remain largely stable over the course of two academic school years.

Accordingly, the most notable difference was not identified across measurement points, but between the two target languages English and French. Motivational dispositions were both higher and more stable for English than for French, where they drop slightly over time. Further research, integrating classroom observational and other qualitative data, could investigate whether and which specific pedagogical dispositions could counteract drops in motivation to learn French as a local foreign language vis-à-vis English as a global lingua franca.

Interestingly, the motivational patterns overall are the same, irrespective of the place of residence: The motivation to learn L2 French is not positively affected when children live close to the French speaking community, as in LAPS I (see Chapter 7). This result would be worth following up from a sociolinguistic perspective, to get a better grasp of the mechanisms that shape the children's rapport with the target languages, these languages' status in the community, the extent of language contact, as well as the need to speak them in order to feel integrated.

To conclude, we are cautiously optimistic that the evidence discussed and presented in this book is a valuable contribution towards a better understanding of individual differences in foreign language learning in a state school setting. We are convinced that the quantitative approach chosen by our group is a possible way to identify generalizable associations between variables. In our take on the topic, we use statistical techniques that should help prevent overfitting of models to data. We used factor analyses (first exploratory, then confirmatory), cross-validation techniques, data partitioning in training and test sets to shed light on the internal dimensionality of our constructs and on the prognostic value of the measures that we took. With very few exceptions, our test instruments were not standardized. Therefore, a direct comparison with other findings in other linguistic contexts is difficult. Also, as argued in Chapter 6, it is easier to estimate the association between constructs if the error with which they have been measured is known. In future endeavours such as ours, it would therefore be important to use more standardized material whose reliability is both known and substantial.

3 Individual differences among scholars in individual difference research

This book is the result of a group effort spanning over roughly five years. Many people who did not co-author chapters have contributed to this project (see acknowledgments in Chapter 0). The chapters of this book are the manifest outcome not only of many hours of joint and individual work, but also of a great

many intensive discussions among the different members of our group. Our group, too, is characterized by individual differences – with respect to research interests, epistemological stances, and language teaching, learning and research experiences. These differences in experience and points of view led to rather lively discussions within the group. Some divergences had to be overcome to arrive at a publishable product with which we all can identify:

Many discussions revolved around questions regarding the selection of the optimal statistical model to test a specific hypothesis, and on the number of models and model variants that should be fitted. The most challenging debates, however, were in the implications of the results of our analyses: What can we infer from the list of variables in the *No costs spared* model in Chapter 4? Are the results in the chapters fitting regression models “good enough” or rather disappointing in terms of variance “explained”? When exactly do associations in structural equation models indeed reveal contribution of one factor to the other and not just association without any further causal meaning? On a more global level, formulating and prioritising the objectives of the project was an ongoing source of discussion – was our main goal to assess the feasibility of prognostication or the analysis of dimensions underlying individual differences? And if both were important, to what extent would such multiple goals compromise the efficiency of the research design? How important is it to do research that provides immediate and applicable answers to practical pedagogical issues? Is it better to measure a multitude of constructs and address a multitude of sub-questions in order to do justice to the “complexity” of multilingual language learning or should we concentrate rigorously on one important research question?

Such discussions are both necessary and difficult. To make the interpretation process even more intricate, some of the manuscript’s reviewers pushed for much bolder conclusions and claims to be drawn from our results than those formulated by us.

Our policy was to discuss these problems repeatedly and extensively. The ultimate decision on the matters pertaining to modelling strategies and to the limits of interpretation, however, were left to the first authors of the chapters respectively. Overall, this approach led to increasingly cautious interpretations of our findings, which may make the read of the book appear overly modest to some. Although the empirical effort might at times seem disproportionate when put in relation to our conclusions, we prefer it that way.

The field of Second Language Acquisition and Multilingualism studies has produced a great wealth of mostly small-scale studies that are often associated with relatively bold and general claims when it comes to the interpretation of the

results. Some such studies from our discipline are (selectively, occasionally) referred to by policy makers, with respect to various aspects of pedagogical policy decisions (curricula, definition of learning outcomes, teaching materials, etc). Research that informs policy must worry about the generalizability of its findings and interpretations. As in other disciplines (cf. Ritchie 2020 for examples mostly from science and psychology), the interpretation of “significant” (or non-significant) results in small-scale studies presents particular and often unrecognized risks for generalizability.

Our group tried to address these problems responsibly by acquiring as much data as was feasible, by collecting two independent data sets from two different contexts and testing patterns emerging from the first with the second, and by carefully cross-validating and testing the models according to the best of our statistical knowledge. To make the results verifiable and future research possible, we publish our data and scripts on the *osf.io* platform, for fellow researchers to download, use and most importantly improve them. We look forward to more discussion with a wider scholarly audience.

References

- Berthele, Raphael. 2019. Policy Recommendations for Language Learning: Linguists’ contributions between scholarly debates and pseudoscience. *Journal of the European Second Language Association* 3(1). 1–11. DOI: 10.22599/jesla.50.
- Conway, Andrew R. A., Michael J. Kane, Michael F. Bunting, D. Zach Hambrick, Oliver Wilhelm & Randall W. Engle. 2005. Working memory span tasks: A methodological review and user’s guide. *Psychonomic Bulletin & Review* 12(5). 769–786. DOI: 10.3758/BF03196772.
- Deci, Edward L. & Richard M. Ryan (eds.). 2002. *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.
- Dörnyei, Zoltán. 2010. *Questionnaires in second language research*. 2nd edn. New York: Routledge.
- Fürst, Guillaume & François Grin. 2018. Multilingualism and creativity: A multivariate approach. *Journal of Multilingual and Multicultural Development* 39(4). 341–355. DOI: 10.1080/01434632.2017.1389948.
- Parry, Thomas S. & Charles W. Stansfield. 1990. *Language aptitude reconsidered*. Englewood, NJ: Prentice Hall Regents.
- Plomin, Robert. 2019. *Blueprint, with a new afterword: How DNA makes us who we are*. London: Penguin Books.
- Ritchie, Stuart. 2020. *Science fictions: How fraud, bias, negligence, and hype undermine the search for truth*. New York: Metropolitan Books.

- Shan, Yiming. 2020. Whether successful language learners require intrinsic motivation. *Open Journal of Modern Linguistics* 10. 549–559. DOI: 10.4236/ojml.2020.105031.
- Skehan, Peter. 2002. Theorising and updating aptitude. In Peter Robinson (ed.), *Individual differences and instructed language learning*, 69–93. Amsterdam/Philadelphia: John Benjamins.
- Stromswold, Karin. 2001. The heritability of language: A review and meta-analysis of twin, adoption, and linkage studies. *Language* 77(4). 647–723. DOI: 10.1353/lan.2001.0247.

Name index

- Abrahamsson, Niclas, 10, 11
Al-Haik, Antoine R., 5
Albert, Ágnes, 126, 128, 137
Alderson, J. Charles, 27
Alexander, Karl L., 32, 107
Alexiou, Thomaï, 14, 15, 201
Allen, Richard J., 62
Ameringer, Victoria, 7
Arnet-Clark, Illya, 130
Au, Shun Y., 29
Austin, Peter C., 133
Avineri, Netta, 31, 109
- Baayen, R. Harald, 149
Babaii, Esmat, 58, 61
Baddeley, Alan D., x, 23, 24
Bader Lehmann, Ursula, 84, 145, 146
Bader, Ursula, 145, 153, 165
Baker, Will, 33
Bandura, Albert, 31
Bates, Douglas, 82, 149, 168
Baumert, Jürgen, vii
Becker, Michael, 84
Berninger, Virginia W., 56, 60, 62
Berthele, Raphael, iv, xi, 33, 63, 122, 211
Bertschy, Ida, 130
Bialystok, Ellen, 12
Biedroń, Adriana, vi, 31, 181
Birdsong, David, 11
Bley-Vroman, Robert, 8
Bollen, Kenneth A., 136
- Bortz, Jürgen, 76
Bourdieu, Pierre, 32, 110, 114
Boyle, Whitney, 23
Breiman, Leo, 97
Brown, Timothy A., 72, 76
Brühwiler, Christian, 145, 146, 152, 156, 166, 174
Brunner, Jerry, 133
Bucher, Monika, 64
Bürkner, Paul-Christian, 205
Busse, Vera, 146, 153, 154, 166
Buyl, Aafke, 146, 153, 166
- Cameron, Lynne, 30
Carless, David R., 33
Carpenter, Helen, 6, 20
Carpenter, Patricia A., 24
Carroll, John B., v, 1–4, 8, 9, 12, 16, 20, 51, 55, 59, 74, 127, 143, 184, 197, 199, 201, 202
Carter, Elaine Fuller, 27
Cattell, Raymond B., 21
Cenoz, Jasone, 122, 164, 165
Chapelle, Carol A., 27
Child, James R., 5
Christiner, Markus, iv, 15
Clément, Richard, 29, 145
Coltheart, Max, 192
Conway, Andrew R. A., 24, 218
Cook, Vivian J., 18, 134
Courtney, Louise, 146, 156
Cropley, Arthur, 26, 127

Name index

- Csapó, Benő, vii
Csizér, Kata, 30, 77, 81, 146, 156
- Dale, Philip S., vi
Daneman, Meredyth, 24
De Bot, Kees, 30
Deci, Edward L., 29, 81, 84, 130, 144,
153, 218
DeKeyser, Robert, 8, 10, 11, 25
Dewaele, Jean-Marc, 146, 165, 175
Dickinson, Leslie, 29
Dörnyei, Zoltán, 26, 29, 30, 57, 75,
126, 144–149, 153, 156, 164,
165, 175, 215
Doughty, Catherine J., 5, 18
Ducroquet, Lucile, 15, 180, 199
- Eckes, Thomas, 58, 183
Edelenbos, Peter, vii
Ehrman, Madeline E., 6
Ellis, Nick C., 30
Ellis, Rod, 28, 33, 54, 126, 129
Elmiger, Daniel, 166, 175
Enever, Janet, vii
Engle, Randall W, 24
Entwisle, Doris R., 32, 107
Erard, Michael, iv
Erlam, Rosemary, 17, 19
Esser, U., 14
Evans, Carol, 27
- Farkas, George, 32, 110
Farsi, Mitra, 27
Fernald, Anne, 107
Field, Andy, 188
Finke, Ronald A., 26, 126, 127
Flege, James, iv
Fleith, Denise de Souza, 128
Foucalt, Michel, 121
- French, Leif M., 23
Friedman, Naomi P., 25
Fröhlich, Maria, 12
Fürst, Guillaume, 219
- Ganschow, Leonore, 8
Gardner, Howard, 21
Gardner, Robert C., 22, 29, 127, 128,
144, 164
Garton, Sue, vii, 33
Gathercole, Susan E., 23, 24
Ghonsooly, Behzad, 126, 128
Giudici, Anja, viii
Gogolin, Ingrid, 32, 120
Golestani, Narly, 8
Graaff, Rick de, 17
Graham, Steve, 56, 62
Graham, Suzanne, 166
Grañena, Gisela, 8, 10, 11, 20, 22, 80
Green, Pat, 27
Grigorenko, Elena L., 3, 16, 19, 26,
127, 199
Grin, François, 219
Grizelj, Sandra, viii
Grotjahn, Rüdiger, 58, 132, 183
Grube, Dietmar, 23, 24
Guilford, Joy Paul, 26, 127
Guthrie, John T., 84
- Häberlin, Urs, 107
Hallquist, Michael, 78
Hansen, Jacqueline, 27
Harber, Kent D., 113
Harrington, Michael, 24, 25
Hasselhorn, Marcus, 23, 24
Heid, Helmut, 121
Heinzmann, Sybille, 31, 57, 75, 81, 145,
146, 149, 152, 153, 156, 164–
166, 174, 175

- Henry, Alastair, 146, 153, 156, 165, 175
Herdina, Philip, 110, 122
Higgins, E. Tory, 30
Hilbert, Sven, 62
Hitch, Graham, x, 23
Holder, Martin C., 146, 156
Horwitz, Elaine, 30, 57, 75, 149
Housen, Alex, 146, 153, 166
Huang, Wenhong, 30, 77, 81
Hufeisen, Britta, 166
Husfeldt, Vera, 84, 145, 146
Hyltenstam, Kenneth, 10, 11

Irwin, Véronique, 109
Iwaniec, Janina, 146

Jaekel, Nils, vii
Jellen, Hans G., 56, 60, 132, 135
Jessner, Ulrike, 27, 110, 122
Johnson, Eric J., 109
Johnson, Jacqueline S., 10
Johnson, Janice, 27
Johnson, Kjell, 94
Johnstone, Richard, vii
Juffs, Alan, 24, 25
Jussim, Lee, 113

Kharkhurin, Anatoliy V., 26, 126, 127
Kigel, Rebecca M., 32
Kiss, Csilla, 11–14, 200, 201
Kline, Rex B., 79, 135
Koeth, Joel, 25
Kolenikov, Stanislav, 136
Kormos, Judit, 6, 16, 19, 24, 30, 77, 81,
128, 137, 146, 200
Kossling, B., 14
Krashen, Stephen, 6, 33, 129
Kreis, Annelies, 84, 145
Kronig, Winfried, 109, 122

Kruidenier, Bastian G., 145
Kübler, Markus, 108
Kuhn, Max, 94
Kuhn, Melanie, 121
Kvalseth, Tarald O., 95

Lambert, Wallace E., 22, 29, 144
Larsen-Freeman, Diane, 30
Le Pape Racine, Christine, 145, 146,
152, 156, 166, 174
Lenhard, Wolfgang, 57, 61, 75, 183
Lenneberg, Eric Heinz, 8
Lenz, Peter, 145, 146
Li, Shaofeng, 6–8, 17, 19, 22, 23, 25,
80, 180, 181, 191, 193
Linck, Jared A., 5, 20, 24, 25
Littlewood, William, 33
Liu, Meihua, 30, 77, 81
Long, Michael H, 10, 11
Lopriore, Lucilla, 165
Lubart, Todd I., 26, 127
Lucas, Samuel R., 109
Lukács, Gabriella, 146

MacIntyre, Peter D., 30, 165
Mackay, Ian R.A., iv
Mai, Miriam, 121
Markus, Hazel, 30
McGroarty, Mary, 164
McNeish, Daniel, 202
Meara, Paul M., 5, 11, 22, 55, 59, 62,
74, 75, 184
Meissner, Franz-Joseph, 166
Mihaljević Djigunović, Jelena, 164,
165
Milton, James, 14, 201
Miyake, Akira, 25, 27
Moïse, Léna Céline, 29
Montanari, Simona, 33, 122

Name index

- Morgan-Short, Kara, 20
Muñoz, Carmen, vii, 12, 14, 201

Newport, Elissa, 10
Nikolov, Marianne, vii, 11, 12, 14, 146,
164, 165, 200
Noels, Kimberly A., 29, 30, 77, 81, 144,
146
Nowicki, Stephen, 57, 59, 75
Nurius, Paula, 30

O'Brien, Irena, 23
O'Hara, Linda A., 127
Oberski, Daniel, 118
Oltman, Philip K., 184
Ottó, István, 11, 12, 126–128, 137, 164

Pajares, Frank, 84
Parry, Thomas S., 5, 215
Pashler, Harold, 20
Pawlak, Mirosław, 181
Peek, Ron, 31
Pekrun, Reinhard, 30, 81
Peng, Peng, 181
Petersen, Calvin R., 5
Peyer, Elisabeth, 57, 75, 145, 146, 149,
152
Pfenninger, Simone, vii, 145, 146, 149,
152, 156, 165
Pickering, Susan J., 23, 24
Pimsleur, Paul, 3, 12, 55, 59, 74, 85,
184, 202
Pinker, Steven, v
Plomin, Robert, 16, 121, 217
Porsch, Raphaela, 58, 61

Quay, Suzanne, 33
Quinn, Terence, 3

Ranta, Leila, 18, 28

Reed, Daniel J., 2, 6, 13
Reiterer, Susanne M., iv, vi, 8, 15, 180
Retelsdorf, Jan, 182
Revelle, William, 202
Ricciardelli, Lina A., 137
Richardson, Julie A., 27
Riener, Cedar, 20
Riley, Dylan, 110
Rimfeld, Kaili, vi, 16, 113
Ritchie, Stuart, 223
Roberts, Cheryl, 27
Roberts, David R., 94
Robinson, Peter, 5, 7, 17, 25
Roehr-Brackin, Karen, 12, 14, 16, 28,
199
Rogers, Vivienne, 5
Rohrer, Julia M., 133
Rosseel, Yves, 78, 111, 116, 133
Rotter, Julian B., 31
Runco, Mark A., 129
Ryan, Richard M., 29, 81, 84, 130, 144,
153, 218
Ryan, Stephen, 26, 126

Sáfár, Anna, 6, 16, 19, 24, 200
Sampasivam, Sinthujaa, 145
Sapon, Stanley M., v, 3, 4, 8, 9, 12, 55,
59, 74, 184, 199, 202
Sasaki, Miyuki, 22, 80
Sattler, Jerome M., 21
Schaer, Ursula, 145, 153, 165
Schaffner, Ellen, 182
Schepens, Job, 32
Schlepppegrell, Mary J., 181
Schmidt, Richard, 29
Schneider, Wolfgang, 57, 61, 75, 183
Schuster, Christof, 76
Shahri, Somaiyeh, 58, 61
Shan, Yiming, 215

- Showqi, Sara, 126, 128
Simonton, Dean K., 132, 137, 138
Singh, Lisa, 166, 175
Singleton, David, 6, 15, 27, 145, 146,
149, 152, 156, 165, 199
Sirin, Selcuk, 182
Skehan, Peter, v, vi, ix, 3, 7, 15, 20, 22,
25, 27, 80, 129, 180, 199, 201,
219
Slobin, Dan I., 106
Snow, Richard, 17
Sparks, Richard L., vi, 8, 31, 81, 181
Spearman, Charles, 21
Speciale, Giovanna, 23
Spolsky, Bernard, 2, 22
St Clair-Thompson, Helen L., 62
Stansfield, Charles W., 2, 6, 13, 27, 215
Sternberg, Robert J., 21, 127
Stöckli, Georg, 30, 57, 75, 77, 81, 145,
146, 149, 152, 153, 165
Stotz, Daniel, viii
Strickland, Bonnie R., 57, 59, 75
Stromswold, Karin, vi, 16, 113, 217
Suárez Vilagran, Maria Del Mar, 11–
14, 16, 200, 201
Sugita McEown, Maya, 146
Taguchi, Tatsuya, 146
Tellier, Angela, 12, 14, 16, 28, 199
Tomasello, Michael, 14
Tragant, Elsa, 165
Turner, Marilyn L., 24
Turner, Tony E., 27
Udry, Isabelle, xi, 33, 63, 122, 181
Urban, Klaus K., 56, 60, 132, 135
Ushioda, Ema, 146–148, 153, 154, 164,
166
Vafaei, Payman, 73
Van der Kloot, Willem A., 78
Van Prooijen, Jan-Willem, 78
Vanhove, Jan, 10, 67, 92, 94, 96, 106,
110, 111, 133
Verspoor, Marjolijn H., 30
Vogt, Thomas, 126, 127
Von Ow, Anna, 145
Wang, Judy Huei-Yu, 84
Waninge, Freya, 30
Weiß, Rudolf H., 56, 60, 62, 74, 76,
132, 135, 184
Wen, Zhisheng (Edward), vi, 7, 16, 20,
24, 25, 181
Wesche, Marjorie, 17, 18, 22, 80
Westfall, Jacob, 94, 133, 135
Wiedenkiller, Eva, 145, 146
Wilden, Eva, 58, 61
William, Revelle, 76, 203
Williams, J. N., 17
Willingham, Daniel, 20
Willis, Jane, 25, 33, 126, 129
Winke, P., 6
Winter, Bodo, 168
Witkin, Herman A., 27, 57, 60, 74
Yaghoubi, Maryam, 27
Yarkoni, Tal, 94, 133, 135
Zemp, Benedict, 64
Zhang, Li-fang, 27

Individual differences in early instructed language learning

Variability in predispositions for language learning has attracted scholarly curiosity for over 100 years. Despite major changes in theoretical explanations and foreign/second language teaching paradigms, some patterns of associations between predispositions and learning outcomes seem timelessly robust. This book discusses evidence from a research project investigating individual differences in a wide variety of domains, ranging from language aptitude over general cognitive abilities to motivational and other affective and social constructs. The focus lies on young learners aged 10 to 12, a less frequently investigated age in aptitude research. The data stem from two samples of multilingual learners in German-speaking Switzerland. The target languages are French and English. The chapters of the book offer two complementary perspectives on the topic: on the one hand, cross-sectional investigations of the underlying structure of these individual differences and their association with the target languages are discussed. Drawing on factor analytical and multivariable analyses, the different components are scrutinized with respect to their mutual dependence and their relative impact on target language skills. The analyses also take into account contextual factors such as the learners' family background and differences across the two contexts investigated. On the other hand, the potential to predict learner's skills in the target language over time based on the many different indicators is investigated using machine learning algorithms. The results provide new insights into the stability of the individual dispositions, on the impact of contextual variables, and on empirically robust dimensions within the array of variables tested.

ISBN 978-3-96110-324-9



9 783961 103249